

# Local convergence and stability of tight bridge-addable graph classes

Guillaume Chapuy<sup>1</sup> ♣ and Guillem Perarnau<sup>2</sup> ♠

♣IRIF, UMR CNRS 8243, *Université Paris-Diderot, France.*

♠CRM, UMI CNRS 3457, *Université de Montréal, Canada.*

♠*School of Mathematics, University of Birmingham, Birmingham, United Kingdom.*

September 14, 2016

## Abstract

A class of graphs is *bridge-addable* if given a graph  $G$  in the class, any graph obtained by adding an edge between two connected components of  $G$  is also in the class. The authors recently proved a conjecture of McDiarmid, Steger, and Welsh stating that if  $\mathcal{G}$  is bridge-addable and  $G_n$  is a uniform  $n$ -vertex graph from  $\mathcal{G}$ , then  $G_n$  is connected with probability at least  $(1 + o_n(1))e^{-1/2}$ . The constant  $e^{-1/2}$  is best possible since it is reached for the class of all forests.

In this paper we prove a form of uniqueness in this statement: if  $\mathcal{G}$  is a bridge-addable class and the random graph  $G_n$  is connected with probability close to  $e^{-1/2}$ , then  $G_n$  is asymptotically close to a uniform  $n$ -vertex random forest in some local sense. For example, if the probability converges to  $e^{-1/2}$ , then  $G_n$  converges in the sense of Benjamini-Schramm to the uniform infinite random forest  $F_\infty$ . This result is reminiscent of so-called “stability results” in extremal graph theory, with the difference that here the stable extremum is not a graph but a graph class.

## 1 Introduction and Main Results

In this paper graphs are simple. A graph is labeled if its vertex set is of the form  $[1..n]$  for some  $n \geq 1$ . An unlabeled graph is an equivalence class of labeled graphs by relabeling. Unless mentioned otherwise, graphs are labeled. A class of (labeled) graphs  $\mathcal{G}$  is *bridge-addable* if given a graph  $G$  in the class, and an edge  $e$  of  $G$  whose endpoints belong to two distinct connected components, then  $G \cup \{e\}$  is also in the class. Examples of bridge-addable classes include planar graphs, forests, or  $H$ -free graphs where  $H$  is any 2-edge connected graph (see many more examples in [ABMR12, CP15]).

McDiarmid, Steger and Welsh [MSW06] conjectured that every bridge-addable class of graphs with  $n$  vertices contains at least a proportion  $(1 + o_n(1))e^{-1/2}$  of connected graphs. This has recently been proved by the authors. In the next statement and later, we denote by  $\mathcal{G}_n$  the set of graphs in  $\mathcal{G}$  with  $n$  vertices, and by  $G_n$  a uniform random element of  $\mathcal{G}_n$ .

---

<sup>1</sup>Support from *Agence Nationale de la Recherche*, grant number ANR 12-JS02-001-01 “Cartapulus”, and from the City of Paris, grant “Émergences Paris 2013, Combinatoire à Paris”. Email: guillaume.chapuy@liafa.univ-paris-diderot.fr.

<sup>2</sup>Email: g.perarnau@bham.ac.uk.

**Theorem A.** [Chapuy, Perarnau [CP15]] For every  $\epsilon > 0$ , there exists an  $n_0$  such that for every bridge-addable class  $\mathcal{G}$  and every  $n \geq n_0$ , we have

$$\Pr(G_n \text{ is connected}) \geq (1 - \epsilon)e^{-1/2}. \quad (1)$$

If  $\mathcal{G}$  is the class of all forests (which is bridge-addable), then Theorem A is asymptotically tight, since it is shown in [Rén59] that if  $F_n$  is a uniform random forest on  $n$  vertices, then as  $n$  goes to infinity:

$$\Pr(F_n \text{ is connected}) \longrightarrow e^{-1/2}. \quad (2)$$

The aim of this paper is to show that, in some appropriate sense, this class is locally the only one for which Theorem A is asymptotically tight. More precisely, we will show that any bridge-addable class of graphs that comes close to achieving the constant  $e^{-1/2}$  is “close” to a uniform random forest in some local sense.

**Definition.** For any  $\zeta > 0$ , we say that a bridge-addable class of graphs  $\mathcal{G}$  is  $\zeta$ -tight with respect to connectivity (or simply  $\zeta$ -tight) if there exists an  $n_0$  such that for every  $n \geq n_0$  we have

$$\Pr(G_n \text{ is connected}) \leq (1 + \zeta)e^{-1/2},$$

where we recall that  $G_n$  is chosen uniformly at random from  $\mathcal{G}_n$ .

In order to state our results, we first need to introduce some notation and terminology. If  $H$  is a graph we let  $|H|$  be its number of vertices. We denote by  $\mathcal{U}$  the set of *unlabeled, unrooted* trees and by  $\mathcal{T}$  the set of *unlabeled, rooted* trees, *i.e.* trees with a marked vertex called the root. For every unrooted tree  $U \in \mathcal{U}$ , we denote by  $\text{Aut}_u(U)$  the number of automorphisms of  $U$ , and for every rooted tree  $T \in \mathcal{T}$ , we denote by  $\text{Aut}_r(T)$  the number of automorphisms of  $T$  that fix its root. Moreover given  $k$  unrooted trees  $U_1, \dots, U_k$  in  $\mathcal{U}$ , we denote by  $\text{Aut}_u(U_1, \dots, U_k)$  the number of automorphisms of the forest formed by disjoint copies of  $U_1, \dots, U_k$ .

Given a graph  $H$ , we let  $\text{Small}(H)$  denote the graph formed by all the components of  $H$  that are not the largest one (in case of a tie, we say that the largest component of the graph is the one with the largest vertex label among all candidates). In what follows, we will always see  $\text{Small}(H)$  as an unlabeled graph. Given a graph  $G$  and a rooted tree  $T \in \mathcal{T}$ , we let  $\alpha^G(T)$  be the number of pendant copies of the tree  $T$  in  $G$ . More precisely,  $\alpha^G(T)$  is the number of vertices  $v$  of  $G$  having the following property: there is at least one cut-edge  $e$  incident to  $v$ , and if we remove the such cut-edge that separates  $v$  from the largest possible component, the vertex  $v$  lies in a component of the graph that is a tree, rooted at  $v$ , which is isomorphic to  $T$ . The following is classical:

**Theorem B** (see Appendix A). *Let  $F_n$  be a uniform random forest with  $n$  vertices. Then for any fixed unlabeled unrooted forest  $\mathbf{f}$  we have as  $n$  goes to infinity:*

$$\Pr(\text{Small}(F_n) \equiv \mathbf{f}) \longrightarrow p_\infty(\mathbf{f}) := e^{-1/2} \frac{e^{-|\mathbf{f}|}}{\text{Aut}_u(\mathbf{f})}, \quad (3)$$

where  $\equiv$  denotes unlabeled graph isomorphism. Moreover,  $p_\infty$  is a probability distribution on the set of unlabeled unrooted forests.

For any fixed rooted tree  $T \in \mathcal{T}$  we have as  $n$  goes to infinity:

$$\frac{\alpha^{F_n}(T)}{n} \xrightarrow{(p)} a_\infty(T) := \frac{e^{-|T|}}{\text{Aut}_r(T)}. \quad (4)$$

where  $(p)$  indicates convergence in probability. Moreover  $a_\infty$  is a probability measure on  $\mathcal{T}$ .

Our main result says that, for a bridge-addable class  $\mathcal{G}$ , if we have an approximate version of (2) for  $G_n$ , then we also have an approximate version of (3) and (4) for  $G_n$ . In the next statement and everywhere in the paper, the constants  $\epsilon, \eta, \rho, \nu, \zeta$  are implicitly assumed (in addition to other written quantifications or assumptions) to be positive and smaller than  $c$  where  $c$  is a small, absolute, constant.

**Theorem 1 (Main result).** *For every  $\epsilon, \eta > 0$ , there exist a  $\zeta > 0$  and an  $n_0$  such that for every  $\zeta$ -tight bridge-addable class  $\mathcal{G}$  and every  $n \geq n_0$ , the following holds:*

- i) The small components of  $G_n$  are close to those of a large random forest, in the sense that for every unrooted unlabeled forest  $\mathbf{f}$  we have:*

$$\left| \Pr(\text{Small}(G_n) \equiv \mathbf{f}) - p_\infty(\mathbf{f}) \right| < \epsilon.$$

- ii) The statistics of pendant trees in  $G_n$  are close to those of a large random forest, in the sense that:*

$$\Pr\left(\forall T \in \mathcal{T} : \left| \frac{\alpha^{G_n}(T)}{n} - a_\infty(T) \right| < \eta\right) > 1 - \epsilon.$$

**Remark 1.** It is easy to see (up to adapting the dependence of  $\zeta$  in  $\epsilon, \eta$ ) that we can replace *i)* by:

- i') The total variation distance between the law of  $\text{Small}(G_n)$  and the probability law  $p_\infty$  is at most  $\epsilon$ .*

Similarly we could replace *ii)* by:

- ii') The  $L^1$ -distance between the measure  $\alpha^{G_n}(\cdot)/n$  and the probability law  $a_\infty(\cdot)$  (both are measures on  $\mathcal{T}$ ) is at most  $\eta$  with probability at least  $1 - \epsilon$ .*

**Remark 2.** Our main result, Theorem 1, can be viewed both as a *unicity result* (since it states that in the limit, and through the lens of local observables, the class of forests is the only one to reach the optimum value  $e^{-1/2}$ ) and as a *stability result* (since it also states that the only classes than come *close* to the extremal value  $e^{-1/2}$  are *close* to forests, again through local observables of random graphs). Here we use the terminology “stability result” on purpose, by analogy with the field of extremal graph theory. Indeed the study of graphs that come *close* to achieving extremal properties is a classical topic in this field. *Stability results*, pioneered in the papers [ES66, Erd67, Erd66, Sim68], show that in many cases, the graphs that are close to being extremal have a structure close to the actual extremal graphs, in some quantifiable sense. Our main result suggests that the question of stability of extremal graph classes, with respect to appropriate graph limit topologies (here, local convergence), should be further examined.

Before going into the proof of the theorem, let us look at some closely related statements and corollaries. Call a bridge-addable class  $\mathcal{G}$  *tight* if it is  $\zeta$ -tight for any  $\zeta > 0$ , that is to say, as  $n$  goes to infinity:

$$\Pr(G_n \text{ is connected}) \rightarrow e^{-1/2}.$$

Then we obtain the following consequence of Theorem 1 for tight bridge-addable classes, which is weaker (it is a unicity, but not a stability result).

**Theorem 2 (Convergence of local statistics in tight bridge-addable graph classes).** *Let  $\mathcal{G}$  be a tight bridge-addable class of graphs. Then, when  $n$  goes to infinity,  $\text{Small}(G_n)$  converges in*

distribution (for the discrete topology) to the probability measure  $p_\infty(\cdot)$  given by (3). Equivalently, for all unrooted forests  $\mathbf{f}$ , we have:

$$\Pr(\text{Small}(G_n) \equiv \mathbf{f}) \longrightarrow p_\infty(\mathbf{f}). \quad (5)$$

Moreover, for any rooted tree  $T \in \mathcal{T}$ , the proportion  $\frac{\alpha^{G_n}(T)}{n}$  of local pendant copies of the tree  $T$  converges in probability to the deterministic constant  $a_\infty(T)$  given by (4):

$$\frac{\alpha^{G_n}(T)}{n} \xrightarrow{(p)} a_\infty(T). \quad (6)$$

In particular, the random measure  $\frac{1}{n}\alpha^{G_n}(\cdot)$  on  $\mathcal{T}$  converges in distribution to the deterministic probability measure  $a_\infty(\cdot)$  (for the topology on measures induced by convergence in distribution on discrete spaces).

Theorem 2 states that, from the point of view of statistics of pendant trees and of non-largest components, tight classes are indistinguishable from random forests in the limit. Let us develop in this direction. Let  $V_n$  be a uniform random vertex in  $G_n$ . Then for a given  $T \in \mathcal{T}$ , conditionally to  $G_n$ , the quantity  $\alpha^{G_n}(T)/n$  is the probability that from  $V_n$  hangs a copy of the tree  $T$ . Readers familiar with the Benjamini-Schramm (BS) convergence of graphs [BS01] will note the similarity with this notion. If  $(G, x)$  and  $(H, y)$  are two rooted graphs (here and in the forthcoming discussion we allow infinite graphs but we always implicitly assume them to be locally finite), define the BS-distance  $d_{BS}((G, x); (H, y))$  as  $2^{-K}$  where  $K$  is the largest integer such that the balls of radius  $K$  in  $(G, x)$  and  $(H, y)$  are isomorphic, as rooted unlabelled graphs. This distance (also called the *ball distance*, see [Lov12]) defines a topology on the set of rooted graphs, and enables us to talk about convergence in distribution of random rooted graphs, in the BS-sense. An equivalent definition of this convergence is the following: a sequence of random rooted graphs  $(H_n, x_n)$  converges to  $(H_\infty, x_\infty)$  if and only if for any  $r \geq 1$  and for any rooted graph  $(H, x)$  of radius  $r$ , the probability that the ball of radius  $r$  in  $(H_n, x_n)$  is isomorphic to  $(H, x)$  converges to the probability of the same event in  $(H_\infty, x_\infty)$ .

It is classical (see e.g. [Ald98, Lemma 2.4]) that if  $F_n$  is a uniform random forest on  $n$  vertices rooted at a random uniform vertex  $V_n$ , then

$$(F_n, V_n) \rightarrow (F_\infty, V_\infty)$$

in distribution in the BS-sense, where  $(F_\infty, V_\infty)$  is the “uniform infinite rooted random forest” (which we could also have called “uniform infinite rooted random tree”, since it is almost surely a tree). Namely,  $(F_\infty, V_\infty)$  can be constructed as follows: consider a semi-infinite path, starting at a vertex  $V_\infty$ , and identify each vertex of this path with the root of an independent Galton-Watson tree with offspring distribution  $\text{Poisson}(1)$ . In our context, passing from pendant trees to balls is an easy task, and one can deduce the following from Theorem 2.

**Corollary 3 (Local convergence of tight bridge-addable graph classes).** *Let  $\mathcal{G}$  be a tight bridge-addable graph class. Let  $G_n$  be a uniform random graph from  $\mathcal{G}_n$  and let  $V_n$  be a uniform random vertex of  $G_n$ . Then  $(G_n, V_n)$  converges to  $(F_\infty, V_\infty)$  in distribution in the Benjamini-Schramm sense.*

The purpose of stating Corollary 3 is to illustrate the link between our local observables and the BS topology, but we could have stated stronger intermediate results. For example Corollary 3 only uses the second part of Theorem 2, and says nothing about the small connected components. In fact, it follows from Theorem 2 that for tight bridge-addable classes, the pair  $((G_n, V_n), \text{Small}(G_n))$  converges in distribution to  $(F_\infty, V_\infty) \otimes p_\infty$  for the product of the BS and the discrete topologies.

Also note that the last corollary is of a weaker nature than Theorem 2 since it averages over the random graph  $G_n$ , while (6) shows that a single sample of the random graph  $G_n$  is locally similar to  $F_\infty$  with high probability. More precisely, it is possible to formulate another, stronger corollary of Theorem 2 in terms of the BS convergence as follows. Let  $\mu_{G_n}$  be the law, given  $G_n$ , of the random rooted graph  $(G_n, V_n)$  where  $V_n$  is a uniform vertex of  $G_n$  (thus  $\mu_{G_n}$  is a random probability measure on the set of locally finite rooted graphs). Then it follows from Theorem 2 that if  $\mathcal{G}$  is a tight bridge-addable class,  $\mu_{G_n}$  converges in distribution to the *deterministic* probability measure  $\mu_\infty$ , defined as the law of  $(F_\infty, V_\infty)$ . The underlying topology for the convergence is the topology of weak convergence induced by the BS distance on the set of probability measures on rooted graphs. We will not give more details on these questions since they only concern weaker reformulations of Theorem 2.

**Remark 3.** Our main theorem asserts that  $\zeta$ -tight bridge-addable classes are “locally similar” to random forests in some precise sense. However, they can be very different from some other perspective. For example, consider the set  $\tilde{\mathcal{F}}_n$  of graphs on  $[1..n]$  defined as follows:  $\tilde{\mathcal{F}}_n$  contains the graph in which all edges linking vertices in  $[1..n^{2/3}]$  are present and all other vertices are isolated, and  $\tilde{\mathcal{F}}_n$  is the smallest bridge-addable class containing this graph. In other words,  $\tilde{\mathcal{F}}_n$  is the set of graphs inducing a clique on  $[1..n^{2/3}]$ , and such that contracting this clique gives a forest. Then  $\tilde{\mathcal{F}} = \cup_{n \geq 1} \tilde{\mathcal{F}}_n$  is a bridge-addable class, and it is easy to see that it is tight (see Appendix A.2 for more details), so our main results apply. However one can argue that the random graph  $\tilde{F}_n$  in  $\tilde{\mathcal{F}}_n$  is *very* different from a random forest in several senses: first, it has  $\Theta(n^{4/3})$  edges whereas a forest has linearly many. Second, with probability  $1 - O(n^{-1/3})$  an edge taken uniformly at random from  $\tilde{F}_n$  belongs to a clique of size  $n^{2/3}$ , which is very different from what happens in a forest. This last point does not contradict our results, but only recalls that it is important here to think of locality as a measure of what happens around “typical vertices” and not “typical edges”.

**Remark 4.** While our results state that typical graphs in tight bridge-addable classes look like random forests in a “local” sense, they do not imply that this similarity is preserved in a “global” sense. For example one can ask whether a random graph in a tight bridge-addable class, properly normalized, converges as a metric space to the Continuous Random Tree (CRT) with respect to the Gromov-Hausdorff (GH) topology. This is true for uniform random forests [Ald93], but a simple example shows that this is not the case in general. Following the lines of the example presented in Remark 3, consider the set  $\hat{\mathcal{F}}_n$  of graphs on  $[1..n]$  defined as follows:  $\hat{\mathcal{F}}_n$  contains the graph where the vertices  $1, \dots, n^{2/3}$  induce a path and all the other ones are isolated, and  $\hat{\mathcal{F}}_n$  is the smallest bridge-addable class containing this graph. Then  $\hat{\mathcal{F}} = \cup_{n \geq 1} \hat{\mathcal{F}}_n$  is a tight bridge-addable class. Nevertheless, the diameter of the random graph  $\hat{F}_n$  is at least  $n^{2/3}$  with probability 1, while the diameter of the largest tree in a uniform random  $n$ -vertex forest is of order  $\sqrt{n}$ . Moreover, when properly renormalized by a scaling factor of  $n^{-2/3}$ ,  $\hat{F}_n$  converges for the GH topology to a real interval and not to the CRT. However, it may be true in general that *typical* distances in tight bridge-addable classes are of order  $\sqrt{n}$ , and even that some convergence to the CRT holds if one allows that a (random) subset of vertices of size  $o(n)$  is removed from the graph. We were not able

to construct a counterexample, and thus leave these questions open.

We conclude this list of results with a simpler statement that does not require the full strength of our main theorems (it is a relatively easy consequence of the results of [CP15], and we will prove it in Section 2).

**Theorem 4.** *Let  $\mathcal{G}$  be a tight bridge-addable class and  $G_n$  a uniform random graph from  $\mathcal{G}_n$ . Then for any  $k \geq 0$ , we have*

$$\Pr(G_n \text{ has } k+1 \text{ connected components}) \longrightarrow e^{-1/2} \frac{2^{-k}}{k!}.$$

*In other words, the number of connected components of  $G_n$  converges in distribution to  $1 + \text{Poisson}(1/2)$ .*

**Structure of the paper.** The proof of Theorem 1 occupies most of the paper, and we now briefly present its structure. The proof roughly follows the one of Theorem A, which we proved in [CP15]: very loosely speaking we show that for a class to be  $\zeta$ -tight, some form of tightness has to occur in each intermediate inequality proven in [CP15]. As the length of the present paper shows, there is however quite an important amount of work to be done to achieve this goal. In particular, apart from some key definitions and notation, and two statements that we directly import from [CP15] (Theorem A and Lemma C), the two papers are disjoint.

We start in Section 2 by proving elementary results about the number of components (including Theorem 4) and we introduce some notions that will play a crucial role in the rest of the proof. Importantly, in Section 2.2, we introduce the partitioning of the space that underlies our technique of local double-counting from [CP15]. In particular we define the notion of “box” that we use in order to partition each graph class according to the local structure of the graphs it contains.

Sections 3 and 4 occupy the most important part of the paper. In Section 3, we prove an analogue of Theorem 1 under the assumption that all elements of  $\mathcal{G}$  are forests. This is done in several steps: In 3.1 we define the notion of “good boxes” and we prove that most of the mass in tight bridge-addable graph classes is localized inside good boxes. These good boxes have the property that they realize locally the extremal value of the optimization problem introduced in [CP15]. This optimization problem expresses some ratios inherited from a double-counting strategy in terms of parameters that record the local structure of the graphs. In 3.2 we study the stability of this problem and deduce that for good boxes, all parameters have to be close to the unique extremum value (closely related to the quantities  $a_\infty$  and  $p_\infty$  appearing in Theorem 1). In 3.3 we use these facts to prove a version of our main result when the graph  $G_n$  has one or two components. Finally in 3.4 we use an induction on the number of components to conclude the proof, in the case of forests.

In Section 4, we address the case of general bridge-addable graph classes: In 4.1 we prove that  $\zeta$ -tight bridge-addable classes tend to have many removable edges (edges that when deleted from a graph in the class, give rise to a graph in the class), and in 4.2 we use this property and the results of Section 3 to conclude the proof of Theorem 1. Finally, in 4.3, we give a detailed proof of Corollary 3.

Appendix A.1 recalls classical results on enumeration of forests and on random forests, while Appendix A.2 gives more details about the example of Remark 3.

For the reader’s convenience we include here a table of contents of the paper:

<b>1</b>	<b>Introduction and Main Results</b>	<b>1</b>
<b>2</b>	<b>First results and set-up for the proof</b>	<b>7</b>
2.1	Number of components in bridge-addable graph classes . . . . .	7
2.2	Partitioning the graph class into highly structured subclasses . . . . .	9
2.3	Notation and quantifiers in the proof . . . . .	10
2.4	Evaluation of generating functions of trees and forests . . . . .	10
<b>3</b>	<b>Theorem 1 for bridge-addable classes of forests</b>	<b>11</b>
3.1	Good and bad boxes . . . . .	11
3.2	Stability of the extremum for the optimization problem . . . . .	14
3.3	Proof of Theorem 1 for classes of forests: the case of 1 or 2 connected components .	19
3.4	Proof of Theorem 1 for classes of forests . . . . .	22
<b>4</b>	<b>From classes of forests to classes of graphs</b>	<b>30</b>
4.1	Removable edges in tight bridge-addable classes of graphs . . . . .	30
4.2	Proof of Theorem 1 . . . . .	35
4.3	Proof of Corollary 3 . . . . .	37
<b>A</b>	<b>Generating functions of trees and forests</b>	<b>41</b>
A.1	Small components and pendant trees in uniform forests . . . . .	41
A.2	More details on the example given in Remark 3 . . . . .	44

## 2 First results and set-up for the proof

In this section, we obtain our first results and we introduce important notions and notation used in the whole paper. In 2.1 we study the number of connected components and we prove Theorem 4. In 2.2, we define the partitioning of the space that underlies our technique of local double-counting. Finally in 2.3, we give a few precisions for the use of quantifiers in the rest of the paper.

### 2.1 Number of components in bridge-addable graph classes

Through the rest of the paper, for a bridge-addable class of graphs  $\mathcal{G}$  and for  $i \geq 1$ , we denote by  $\mathcal{G}_n^{(i)}$  the set of  $n$ -vertex graphs in  $\mathcal{G}$  having  $i$  connected components. An elegant double-counting argument going back to [MSW06] asserts that for all  $i \geq 1$ , and  $n \geq 1$  we have:

$$i \cdot \left| \mathcal{G}_n^{(i+1)} \right| \leq \left| \mathcal{G}_n^{(i)} \right|. \quad (7)$$

This statement follows by double-counting the edges of an auxiliary bipartite graph on the vertex set  $\mathcal{G}_n^{(i)} \uplus \mathcal{G}_n^{(i+1)}$ , where two graphs  $G, H$  are linked by an edge if and only if one can be obtained from the other by adding a bridge: on the one hand, an element of  $\mathcal{G}_n^{(i+1)}$  has degree at least  $i(n-i)$  in this auxiliary graph, since  $\mathcal{G}$  is bridge-addable; on the other hand, an element of  $\mathcal{G}_n^{(i)}$  has degree at most  $(n-i)$  (which is the maximum number of cut-edges in a graph with  $i$  connected components and  $n$  vertices). Thus (7) follows. The main achievement of the paper [CP15] was to improve this bound by roughly a factor  $\frac{1}{2}$ , asymptotically.

**Lemma C** ([CP15, Proposition 5]). *For every  $\eta$  and every  $m$ , if  $\mathcal{G}$  is a bridge-addable class and  $n$  is large enough, one has for every  $i \leq m$ ,*

$$i|\mathcal{G}_n^{(i+1)}| \leq \left(\frac{1}{2} + \eta\right) |\mathcal{G}_n^{(i)}|. \quad (8)$$

The following lemma, which follows relatively easily from Lemma C, provides a converse inequality to (8) for  $\zeta$ -tight classes. Note that it implies Theorem 4.

**Lemma 5.** *For every  $\eta$  and every  $m$  there exists a  $\zeta$  such that for every  $\zeta$ -tight bridge-addable class  $\mathcal{G}$  and provided  $n$  is large enough, one has for every  $i \leq m$ ,*

$$\left(\frac{1}{2} - \eta\right) |\mathcal{G}_n^{(i)}| \leq i|\mathcal{G}_n^{(i+1)}| \leq \left(\frac{1}{2} + \eta\right) |\mathcal{G}_n^{(i)}|.$$

*Proof.* The second inequality is precisely Lemma C.

The path to prove the first inequality is rather straightforward, but we will proceed carefully since the order in which the different parameters are chosen requires some caution. We proceed by contradiction. Fix  $\eta$  and  $m$  and assume that for every  $\zeta > 0$  there exist a  $\zeta$ -tight bridge-addable class  $\mathcal{G}$ , a large enough  $n_*$  and an  $i_* \leq m$  such that

$$i_*|\mathcal{G}_{n_*}^{(i_*+1)}| \leq \left(\frac{1}{2} - \eta\right) |\mathcal{G}_{n_*}^{(i_*)}|. \quad (9)$$

Let  $i_0 \geq m$  be an integer that we will choose later. By Lemma C, if  $n$  is large enough, (8) holds with  $\eta = \zeta$  for any  $i \leq i_0$ . Also, since  $\mathcal{G}$  is  $\zeta$ -tight, provided that  $n_*$  is large enough, we have

$$\frac{|\mathcal{G}_{n_*}^{(1)}|}{|\mathcal{G}_{n_*}|} \leq (1 + \zeta)e^{-1/2}. \quad (10)$$

Noting  $f_i(x) := \sum_{j>i} \frac{x^j}{j!}$ , we can now bound the inverse of the probability that  $G_{n_*}$  is connected as follows

$$\begin{aligned} \frac{|\mathcal{G}_{n_*}|}{|\mathcal{G}_{n_*}^{(1)}|} &\leq \sum_{i=1}^{i_*-1} \frac{|\mathcal{G}_{n_*}^{(i)}|}{|\mathcal{G}_{n_*}^{(1)}|} + \sum_{i=i_*}^{i_0} \frac{|\mathcal{G}_{n_*}^{(i)}|}{|\mathcal{G}_{n_*}^{(1)}|} + \sum_{i \geq i_0+1} \frac{|\mathcal{G}_{n_*}^{(i)}|}{|\mathcal{G}_{n_*}^{(1)}|} \\ &\leq \sum_{i=1}^{i_*-1} \frac{1}{i!} \left(\frac{1}{2} + \zeta\right)^i + \sum_{i=i_*}^{i_0} \frac{1}{i!} \left(\frac{1}{2} + \zeta\right)^i \frac{\frac{1}{2} - \eta}{\frac{1}{2} + \zeta} + f_{i_0}(1) \end{aligned}$$

where for the last term we used the bound (7). Thus:

$$\begin{aligned} \frac{|\mathcal{G}_{n_*}|}{|\mathcal{G}_{n_*}^{(1)}|} &\leq e^{\frac{1}{2} + \zeta} - f_{i_0}(1/2 + \zeta) + \left(\frac{\frac{1}{2} - \eta}{\frac{1}{2} + \zeta} - 1\right) (f_{i_*-1}(1/2 + \zeta) - f_{i_0}(1/2 + \zeta)) + f_{i_0}(1) \\ &\leq e^{\frac{1}{2} + \zeta} - \frac{\eta + \zeta}{1/2 + \zeta} \cdot f_{i_*-1}(1/2) + f_{i_0}(1) \\ &\leq e^{1/2} + (e^\zeta - 1)e^{1/2} - \eta f_m(1/2) + f_{i_0}(1) \end{aligned}$$



We now choose  $\zeta$  small enough with respect to  $\eta$  and  $m$  such that  $\frac{\eta}{2}f_m(1/2) \geq (e^\zeta - 1 + 2\zeta)e^{1/2}$ , and we choose  $i_0$  large enough with respect to  $m$ , in such a way that  $\frac{\eta}{2}f_m(1/2) \geq f_{i_0}(1)$ . These choices fix the value  $n_*$  as above, and we finally get the bound:

$$\frac{|\mathcal{G}_{n_*}^{(1)}|}{|\mathcal{G}_{n_*}|} \geq (1 - 2\zeta)^{-1}e^{-1/2} \geq (1 + 2\zeta)e^{-1/2},$$

However, since  $n_*$  is arbitrarily large, we obtain a contradiction with (10).  $\square$

## 2.2 Partitioning the graph class into highly structured subclasses

We now introduce a partitioning of  $\mathcal{G}_n$  in terms of some local statistics, which requires the following set-up, that is modeled on [CP15, proof of Prop 3]. Here  $\epsilon$  and  $k_*$  are two constants, whose value may vary along the course of the paper, that will *in fine* be chosen very small and very large, respectively:

- $\mathcal{U}_\epsilon$  is the set of unrooted trees of order at most  $\lceil \epsilon^{-1} \rceil$ :

$$\mathcal{U}_\epsilon := \{U \in \mathcal{U}, |U| \leq \lceil \epsilon^{-1} \rceil\}.$$

- $\mathcal{T}_*$  is the set of rooted trees of order at most  $k_*$ :

$$\mathcal{T}_* := \{T \in \mathcal{T}, |T| \leq k_*\}.$$

More generally, for any given  $\ell \geq 1$ , we will use the notation  $\mathcal{T}_{\leq \ell}$  (resp.,  $\mathcal{U}_{\leq \ell}$ ) to denote the set of rooted (resp., unrooted) trees of order at most  $\ell$ . Then,  $\mathcal{U}_\epsilon = \mathcal{U}_{\leq \lceil \epsilon^{-1} \rceil}$  and  $\mathcal{T}_* = \mathcal{T}_{\leq k_*}$ .

Roughly speaking, we will use elements of  $\mathcal{U}_\epsilon$  and  $\mathcal{T}_*$  as “test graphs” to measure the shape of small components of  $G_n$  and the number of pending subtrees of  $G_n$  of given shapes, respectively. For  $\ell \geq 1$ , we write  $\mathcal{E}_\ell = \{0, \dots, n-1\}^{\mathcal{T}_{\leq \ell}}$ , and we will be particularly concerned with the set  $\mathcal{E}_{k_*}$ , namely the set of integer vectors with one coordinate for each “test tree” in  $\mathcal{T}_*$ . For  $\alpha \in \mathcal{E}_{k_*}$  and  $w \geq 1$  (width), we define the *box*  $[\alpha]^w \subset \mathcal{E}_{k_*}$  and its *q-neighborhood*  $[\alpha]_q^w$  as the parallelepipeds:

$$\begin{aligned} [\alpha]^w &:= \{\alpha' \in \mathcal{E}_{k_*} : \forall T \in \mathcal{T}_*, \alpha(T) \leq \alpha'(T) < \alpha(T) + w\}, \\ [\alpha]_q^w &:= \{\alpha' \in \mathcal{E}_{k_*} : \forall T \in \mathcal{T}_*, \alpha(T) - q \leq \alpha'(T) < \alpha(T) + w + q\}. \end{aligned}$$

Note that here, and elsewhere in the paper, we slightly abuse notation by using both the letter  $\alpha$  to denote an element of  $\mathcal{E}_\ell$  and the notation  $\alpha^G$  to denote the function  $\alpha^G : \mathcal{T} \rightarrow \mathcal{E}_\ell$  that counts the number of pendant trees of a given shape in the graph  $G$ .

Finally, if  $\mathcal{S}_n$  denotes a set of graphs (where the letter  $\mathcal{S}$  could carry other decorations), we let  $\mathcal{S}_{n,[\alpha]^w}$  be the set of graphs  $G$  in  $\mathcal{S}_n$  such that  $(\alpha^G(T))_{T \in \mathcal{T}_*} \in [\alpha]^w$ , and we use the same notation with  $[\alpha]_q^w$ .

Also, for every forest  $\{U_1, \dots, U_k\}$ , we denote by  $\mathcal{S}_n^{\{U_1, \dots, U_k\}}$  the set of graphs  $G$  in  $\mathcal{S}_n$  such that  $\text{Small}(G)$  is isomorphic to  $\{U_1, \dots, U_k\}$ . While we denote a forest by  $\{U_1, \dots, U_k\}$ , one should understand it as an *unordered multiset* of unrooted trees. In the case of graphs with two connected components, we just use the notation  $\mathcal{S}_n^U$  for  $\mathcal{S}_n^{\{U\}}$ , where  $U \in \mathcal{U}$ .

### 2.3 Notation and quantifiers in the proof

The proof of Theorem 1 involves many different quantifiers. In the statements and everywhere in the paper, all Greek letters, apart from  $\alpha$  and  $\beta$ , are implicitly assumed (in addition to other written quantifications or assumptions) to be positive constants that are smaller than  $c$  where  $c$  is a small, absolute, constant. The letters  $\alpha$  and  $\beta$  are used to refer to elements of the space  $\mathcal{E}_{k*}$  or  $\mathcal{E}_\ell$ . We also use latin letters to denote integers that are greater than or equal to one.

Each statement in Sections 3 and 4 involves several variables and the relative dependency between them plays a subtle role in the proof. We have carefully made all quantifiers explicit in all the statements. However, the reader can use the following inequalities to clarify the hierarchy of (small) parameters used in Sections 3 and 4:

$$\frac{1}{n} \ll \zeta \ll \frac{1}{w} \ll \frac{1}{k_*} \ll \xi \ll \epsilon = \frac{1}{q} \ll \gamma \ll \rho \ll \nu \ll \vartheta \ll \eta \ll \theta_1 \ll \dots \ll \theta_k, \delta, 1/\ell, 1/k, 1/u \leq 1, \quad (11)$$

where the notation  $a \ll b \leq 1$  has to be read as: In each statement involving both variables  $a$  and  $b$ , there exists a non-decreasing function  $f : (0, 1] \rightarrow (0, 1]$  such that the statement holds for every  $0 < a \leq b \leq 1$  such that  $a \leq f(b)$ . For example, the order in which the quantifiers appear in the statement of Lemma 5 above correspond to the notation:

$$n^{-1} \ll \zeta \ll \eta, m^{-1}.$$

Note that  $1/n$  is the leftmost quantity appearing in (11): throughout the paper,  $n$  will be taken arbitrarily large with respect to all the other constants. We will often write “for  $n$  large enough” rather than “there exists  $n_0$  such that for  $n \geq n_0$ ...”, and it will always be the case that the implicit value of the parameter  $n_0$  is chosen after all other parameters, and may depend on all of them.

During the proof, we will use the notation  $a = b \pm \mu$  to denote that  $b - \mu \leq a \leq b + \mu$ .

### 2.4 Evaluation of generating functions of trees and forests

In this subsection we recall two classical evaluations of generating functions of trees and forests that we will use several times in our proofs. We let  $T(z) = \sum_{n \geq 1} \frac{t_n}{n!} z^n$  be the exponential generating function of all rooted labelled trees, where the exponent of the variable  $z$  records the number of vertices. Hence  $t_n$  is the number of rooted trees on  $[1..n]$ , given by Cayley’s formula:  $t_n = n^{n-1}$ . We also let  $F(z) = \sum_{n \geq 0} \frac{f_n}{n!} z^n$  be the exponential generating function of (unrooted) labelled forests (here  $f_n$  is the number of forests on the vertex set  $[1..n]$ , and by convention  $f_0 = 1$ ). We have:

**Lemma 6.** *Both  $T(z)$  and  $F(z)$  have radius of convergence  $e^{-1}$ , and both are finite at their main singularity  $z = e^{-1}$ , where we have  $T(e^{-1}) = 1$  and  $F(e^{-1}) = e^{1/2}$ . Moreover for  $z$  in a slit neighbourhood of  $e^{-1}$  we have*

$$T(z) = 1 + O(\sqrt{1 - ze}) . \quad (12)$$

The proof is a classical exercise in analytic combinatorics and is recalled in Appendix A.

### 3 Theorem 1 for bridge-addable classes of forests

Balister, Bollobás and Gerke [BBG08, Lemma 2.1] proposed an elegant argument that reduces the statement of Theorem A to the case where all graphs in  $\mathcal{G}$  are forests. As we will see in the next section, their idea can be adapted to the present context. We will therefore start by proving Theorem 1 for classes  $\mathcal{G}$  composed by forests:

*Throughout the rest of Section 3, we will assume that all graphs in  $\mathcal{G}$  are forests.*

#### 3.1 Good and bad boxes

The main concern of the paper [CP15] was to obtain a version of the double-counting argument of Section 2.1 that is *local* in the sense that it relates cardinalities of graphs corresponding to fixed boxes.

In order to select a collection of boxes, we will focus on the graphs in  $\mathcal{G}_n$  that have either one or two connected components, and, in view of this, we use the shorter notation  $\mathcal{A}_n := \mathcal{G}_n^{(1)}$  and  $\mathcal{B}_n := \mathcal{G}_n^{(2)}$ .

Given  $\epsilon$  (hence  $\mathcal{U}_\epsilon$ ) and  $k_*$  (hence  $\mathcal{T}_*$ ), [CP15, Lemma 17] asserts that there exist integers  $K$  and  $w$  (independent of  $\mathcal{G}$  and of  $n$ ) and a set of  $K$  disjoint boxes of width  $w$  in  $\mathcal{E}_{k_*}$ , noted  $\{[\beta_i]^w, 1 \leq i \leq K\}$ , such that if  $q = q_\epsilon := \lceil \epsilon^{-1} \rceil$  and if  $n$  is large enough, then the  $q$ -neighbourhoods of boxes form a partition of  $\mathcal{E}_{k_*}$ ,

$$\biguplus_{i=1}^K [\beta_i]_q^w = \mathcal{E}_{k_*}, \quad (13)$$

and moreover for each  $U \in \mathcal{U}_\epsilon$  we have:

$$\sum_{i=1}^K |\mathcal{B}_{n, [\beta_i]^w}^U| \geq (1 - \epsilon) |\mathcal{B}_n^U|. \quad (14)$$

Note that from (13), the boxes  $[\beta_i]^w$  are  $2q$ -apart from each other, and yet (14) ensures that they capture a proportion at least  $(1 - \epsilon)$  of the set  $\mathcal{B}_n^U$  for each  $U \in \mathcal{U}_\epsilon$ . We now fix such a set of boxes, and we will keep referring to these boxes  $([\beta_i]^w)_{1 \leq i \leq K}$  (or simply,  $([\beta_i])_{1 \leq i \leq K}$ ) throughout Section 3, keeping in mind that the number  $K = K(\epsilon, k_*)$  of boxes, and their width  $w = w(\epsilon, k_*)$ , depend on  $\epsilon$  and  $k_*$  but neither on  $\mathcal{G}$  nor on  $n$ .

In the present paper, one of the main tasks consists in showing that the global estimates obtained in [CP15], such as Lemma C, can be “lowered” down to boxes for  $\zeta$ -tight classes. This is not true for every box in  $\mathcal{E}_{k_*}$ , but it will be for certain boxes that contain most of the graphs in the class. For every  $\gamma$  and every  $\epsilon$ , we say that a box  $[\alpha]^w$  is  $(\gamma, \epsilon)$ -*good* (or simply *good*) if the two following conditions hold:

- i)  $|\mathcal{B}_{n, [\alpha]^w}| \geq (\frac{1}{2} - \gamma) \cdot |\mathcal{A}_{n, [\alpha]^w}|$ , and
- ii)  $\sum_{U \notin \mathcal{U}_\epsilon} |\mathcal{B}_{n, [\alpha]^w}^U| < \gamma |\mathcal{B}_{n, [\alpha]^w}|$ .

Note that Property i) is a local version of the first inequality of Lemma 5 for  $i = 1$ , while Property ii) ensures that the number of graphs in sets that we do not control, is small. Hence good boxes are, in some sense, boxes that realize the tightness property locally.

We will be interested in the boxes among the  $[\beta_i]$  that are  $(\gamma, \epsilon)$ -good:

$$\text{Good}_{\gamma, \epsilon} := \{i \in [1..K] : [\beta_i] \text{ is } (\gamma, \epsilon)\text{-good}\} .$$

An important step in the proof of Theorem 1 is the following result:

**Lemma 7.** *For every  $\gamma$  and every  $\eta$ , if  $\epsilon < \epsilon_0(\gamma, \eta)$  and if  $k_* \geq k_0(\epsilon)$ , then there exists  $\zeta$  such that for every  $\zeta$ -tight bridge-addable class  $\mathcal{G}$  and every large enough  $n$ , one has*

$$\frac{\sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]_q^w}|}{|\mathcal{A}_n|} < \eta ,$$

and

$$\frac{\sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{B}_{n, [\beta_i]^w}|}{|\mathcal{B}_n|} < \eta .$$

*Proof.* Let  $\epsilon > 0$  (to be fixed later). Up to setting  $k_*$  and  $n$  large enough, we can use Equation (16) in [CP15] for each  $1 \leq i \leq K$ ,

$$\sum_{U \in \mathcal{U}_\epsilon} |\mathcal{B}_{n, [\beta_i]^w}^U| \leq \frac{1}{2} \cdot |\mathcal{A}_{n, [\beta_i]_q^w}| (1 + 3\epsilon) \leq \left(\frac{1}{2} + 2\epsilon\right) \cdot |\mathcal{A}_{n, [\beta_i]_q^w}| .$$

Moreover, provided that  $n$  is large enough, we have (Equation (17) in [CP15])

$$\sum_{U \notin \mathcal{U}_\epsilon} |\mathcal{B}_n^U| \leq 2\epsilon |\mathcal{A}_n| \tag{15}$$

From the last two inequalities, we have

$$\sum_{i \in \text{Good}_{\gamma, \epsilon}} |\mathcal{B}_{n, [\beta_i]^w}| \leq 2\epsilon |\mathcal{A}_n| + \left(\frac{1}{2} + 2\epsilon\right) \sum_{i \in \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]_q^w}| . \tag{16}$$

Let  $S$  and  $T$  be the sets of indices  $i \notin \text{Good}_{\gamma, \epsilon}$  such that  $[\beta_i]^w$  violates i) and ii) respectively. Using (15), we have

$$\sum_{i \in T} |\mathcal{B}_{n, [\beta_i]^w}| \leq \sum_{i \in T} \frac{1}{\gamma} \sum_{U \notin \mathcal{U}_\epsilon} |\mathcal{B}_{n, [\beta_i]^w}^U| \leq \frac{1}{\gamma} \sum_{U \notin \mathcal{U}_\epsilon} |\mathcal{B}_n^U| \leq \frac{2\epsilon}{\gamma} |\mathcal{A}_n| .$$

From the previous equation it follows that

$$\begin{aligned} \sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{B}_{n, [\beta_i]^w}| &\leq \sum_{i \in S} |\mathcal{B}_{n, [\beta_i]^w}| + \sum_{i \in T} |\mathcal{B}_{n, [\beta_i]_q^w}| \\ &\leq \left(\frac{1}{2} - \gamma\right) \sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]_q^w}| + \frac{2\epsilon}{\gamma} |\mathcal{A}_n| . \end{aligned} \tag{17}$$

Using (16) and (17), we get

$$\begin{aligned}
(\gamma + 2\epsilon) \sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]_q^w}| &\leq (\gamma + 2\epsilon) \sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]_q^w}| + \sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{B}_{n, [\beta_i]_q^w}| \\
&+ \sum_{i \in \text{Good}_{\gamma, \epsilon}} |\mathcal{B}_{n, [\beta_i]_q^w}| - \sum_{i=1}^K |\mathcal{B}_{n, [\beta_i]_q^w}| \\
&\leq (\gamma + 2\epsilon) \sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]_q^w}| + \left(\frac{1}{2} - \gamma\right) \sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]_q^w}| \\
&+ \left(\frac{1}{2} + 2\epsilon\right) \sum_{i \in \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]_q^w}| + \frac{4\epsilon}{\gamma} |\mathcal{A}_n| - \sum_{i=1}^K |\mathcal{B}_{n, [\beta_i]_q^w}|
\end{aligned}$$

The last inequality can be simplified as

$$\begin{aligned}
(\gamma + 2\epsilon) \sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]_q^w}| &\leq \left(\frac{1}{2} + 2\epsilon\right) \sum_{i=1}^K |\mathcal{A}_{n, [\beta_i]_q^w}| - \sum_{i=1}^K |\mathcal{B}_{n, [\beta_i]_q^w}| + \frac{4\epsilon}{\gamma} |\mathcal{A}_n| \\
&\leq \left(\frac{1}{2} + \frac{6\epsilon}{\gamma}\right) |\mathcal{A}_n| - \sum_{i=1}^K |\mathcal{B}_{n, [\beta_i]_q^w}|
\end{aligned} \tag{18}$$

where we used that the  $[\beta_i]_q^w$  are disjoint. Using (15) again and (14), we have

$$\begin{aligned}
\sum_{i=1}^K |\mathcal{B}_{n, [\beta_i]_q^w}| &\geq \sum_{i=1}^K \sum_{U \in \mathcal{U}_\epsilon} |\mathcal{B}_{n, [\beta_i]_q^w}^U| \\
&\geq (1 - \epsilon)(|\mathcal{B}_n| - 2\epsilon |\mathcal{A}_n|) \\
&\geq (1 - \epsilon)|\mathcal{B}_n| - 2\epsilon |\mathcal{A}_n|
\end{aligned}$$

Finally, Lemma 5 with  $i = 1$  and  $\eta$  replaced by  $\epsilon$ , implies that if  $\zeta$  is small enough,  $\mathcal{G}$  is  $\zeta$ -tight and  $n$  is large enough, then the last quantity is larger than  $(1/2 - 4\epsilon) |\mathcal{A}_n|$ .

We now choose  $\epsilon_0 = \frac{\eta}{20}$ . Going back to (18), if  $\epsilon < \epsilon_0$ , we get

$$\sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]_q^w}| \leq \frac{10\epsilon}{\gamma(\gamma + 2\epsilon)} |\mathcal{A}_n| \leq \frac{\eta}{2} |\mathcal{A}_n|, \tag{19}$$

which proves the first part of the lemma.

For the second part of the lemma, we use (17) and Lemma 5 with  $\eta$  replaced by  $\epsilon$ , to get

$$\frac{\sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{B}_{n, [\beta_i]_q^w}|}{|\mathcal{B}_n|} \leq \frac{\left(\frac{1}{2} - \gamma\right) \sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]_q^w}| + \frac{2\epsilon}{\gamma} |\mathcal{A}_n|}{\left(\frac{1}{2} - \epsilon\right) |\mathcal{A}_n|}$$

By (19), we conclude

$$\frac{\sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{B}_{n, [\beta_i]_q^w}|}{|\mathcal{B}_n|} \leq \frac{\left(\frac{1}{2} - \gamma\right) \frac{\eta}{2} |\mathcal{A}_n| + \frac{2\epsilon}{\gamma} |\mathcal{A}_n|}{\left(\frac{1}{2} - \epsilon\right) |\mathcal{A}_n|} \leq \eta. \quad \square$$

### 3.2 Stability of the extremum for the optimization problem

The goal of this subsection is to estimate the ratio between  $|\mathcal{B}_{n,[\alpha]^w}^U|$  and  $|\mathcal{A}_{n,[\alpha]^w}|$ , when  $[\alpha]^w$  is a good box and  $U \in \mathcal{U}_\epsilon$ . Recall that good boxes are the ones that locally inherit the global property of being tight; for instance, in good boxes we have a precise estimation of the ratio between  $|\mathcal{B}_{n,[\alpha]^w}|$  and  $|\mathcal{A}_{n,[\alpha]^w}|$ .

In order to do that, we will need to return to the original “optimization problem” introduced in [CP15]. Namely, we will study certain functionals of the ratios  $|\mathcal{B}_{n,[\alpha]^w}^U|/|\mathcal{A}_{n,[\alpha]^w}|$ , or more precisely of the variables  $(z_{n,\alpha}^U)_{U \in \mathcal{U}_\epsilon}$ , defined by (20) below. We will proceed as follows: Lemma 8 gives the “constraints” of the optimization problem, by showing that the variables  $z_{n,\alpha}^U$  have to be close to a certain domain  $D$ ; Lemma 9 shows that if  $[\alpha]^w$  is good, then the “objective function” of the optimization problem has to be close to its optimal value given these constraints (which was proved to be  $\frac{1}{2}$  in [CP15]). Then Lemma 10 proves a form of unicity of the extremum. From these three lemmas we deduce the main results of this subsection: if  $[\alpha]^w$  is good, then  $(z_{n,\alpha}^U)_{U \in \mathcal{U}_\epsilon}$  is close to  $p_\infty(U)$  for each unrooted tree  $U$  of bounded size (Proposition 11) and if  $[\alpha]^w$  is good, then  $\alpha(T)/n$  is close to  $a_\infty(T)$  for each rooted tree  $T$  of bounded size (Proposition 12).

Apart from the proof of Lemma 7 already given, the proofs of Lemmas 8–9–10 are the part of the present paper that rely the most on [CP15]. Indeed, we will refer to several technical statements therein in our proofs. This will no longer be the case in the next sections.

Following [CP15], given  $\epsilon$  (hence  $\mathcal{U}_\epsilon$ ) we define a  $\mathcal{U}_\epsilon$ -admissible decomposition of  $T$  as an increasing sequence  $\mathbf{T} = (T_i)_{i \leq \ell}$  of labeled trees

$$T_1 \subset \cdots \subset T_\ell = T$$

for some  $\ell \geq 1$  called the *length*, such that  $T_1 \in \mathcal{U}_\epsilon$  and, for each  $2 \leq i \leq \ell$ ,  $T_i$  is obtained by joining  $T_{i-1}$  by an edge  $e_i$  to some tree  $U_i \in \mathcal{U}_\epsilon$ . The *weight* of  $\mathbf{T}$  with respect to  $\mathbf{z} = (z^U)_{U \in \mathcal{U}_\epsilon} \in (\mathbb{R}_+)^{\mathcal{U}_\epsilon}$  is defined as  $\omega(\mathbf{T}, \mathbf{z}) = \prod_{i=1}^\ell z^{U_i}$ , where  $U_i = T_i \setminus T_{i-1}$  as an unrooted tree (here we use the convention  $T_0 = \emptyset$ ). The *maximum weight* of  $T$  with respect to  $\mathbf{z}$ , denoted by  $\omega(T, \mathbf{z})$ , is defined as the maximum of  $\omega(\mathbf{T}, \mathbf{z})$  over all the  $\mathcal{U}_\epsilon$ -admissible decompositions  $\mathbf{T}$  of  $T$ .

We now use  $\omega(T, \mathbf{z})$  to define the following partition functions,

$$\begin{aligned} Y(\mathbf{z}) &:= \sum_{T \in \mathcal{T}} \frac{\omega(T, \mathbf{z})}{\text{Aut}_r(T)}, & Y^u(\mathbf{z}) &:= \sum_{U \in \mathcal{U}} \frac{\omega(U, \mathbf{z})}{\text{Aut}_u(U)}, \\ Y_{\mathcal{T}_*}(\mathbf{z}) &:= \sum_{T \in \mathcal{T}_*} \frac{\omega(T, \mathbf{z})}{\text{Aut}_r(T)}, & Y_{\mathcal{U}_\epsilon}^u(\mathbf{z}) &:= \sum_{U \in \mathcal{U}_\epsilon} \frac{\omega(U, \mathbf{z})}{\text{Aut}_u(U)} \\ Y_{\leq k}(\mathbf{z}) &:= \sum_{T \in \mathcal{T}_{\leq k}} \frac{\omega(T, \mathbf{z})}{\text{Aut}_r(T)}, & \tilde{Y}_{\mathcal{U}_\epsilon}^u(\mathbf{z}) &:= \sum_{U \in \mathcal{U}_\epsilon} \frac{z^U}{\text{Aut}_u(U)}. \end{aligned}$$

Furthermore, we define the domain of convergence of  $Y(\mathbf{z})$  as follows,

$$D := \{\mathbf{z} \in (\mathbb{R}_+)^{\mathcal{U}_\epsilon}, Y(\mathbf{z}) < \infty\}.$$

It is important to note that there is an implicit dependence of  $\omega(T, \mathbf{z})$  on  $\epsilon$  (via  $\mathcal{U}_\epsilon$ -admissible decompositions). Hence, all the partition functions defined above (and their respective domains) also depend on  $\epsilon$ . In order to keep the notation light we do not make this dependence explicit.

Let  $\mathbf{j} := (1)_{U \in \mathcal{U}_\epsilon}$  be the all-one vector of length  $|\mathcal{U}_\epsilon|$ . Given a choice of  $n$ , to each  $\alpha \in \mathcal{E}_{k_*}$  we assign a vector  $\mathbf{z}_{n,\alpha} = (z_{n,\alpha}^U)_{U \in \mathcal{U}_\epsilon} \in (\mathbb{R}_+)^{\mathcal{U}_\epsilon}$ , where

$$z_{n,\alpha}^U := \text{Aut}_u(U) \frac{|\mathcal{B}_{n,[\alpha]}^U|}{|\mathcal{A}_{n,[\alpha]}^w|} \left(1 - \frac{|U|}{n}\right), \quad (20)$$

where  $q = \lceil \epsilon^{-1} \rceil$  as before and  $w = w(\epsilon, k_*)$  is chosen as in Section 3.1.

**Lemma 8.** *For every  $\xi$  and every  $\epsilon$ , if  $k_* \geq k_0(\epsilon, \xi)$  and  $n$  is large enough, then for every  $\alpha \in \mathcal{E}_{k_*}$  we have that  $\mathbf{z}_{n,\alpha} - \xi \mathbf{j} \in D$ .*

*Proof.* For the sake of contradiction, assume that there exist  $\xi$  and  $\epsilon$  such that for every  $k_0$  there exists  $k \geq k_0$  such that for every large enough  $n$  there exists  $\alpha_{n,k} \in \mathcal{E}_k$  with

$$\mathbf{z}_{n,\alpha_{n,k}} - \xi \mathbf{j} \notin D.$$

For a given  $k \geq k_0$ , let  $\mathbf{z}_k$  be a limit point of the sequence  $(\mathbf{z}_{n,\alpha_{n,k}})_{n \geq 1}$ . Since  $D$  is closed downwards (Lemma 13 in [CP15]), then  $\mathbf{z}_k - \xi \mathbf{j} \notin D$ .

Moreover, by Corollary 12 in [CP15], we have  $Y_{\leq k}(\mathbf{z}_k) \leq 1$ . Similarly as in [CP15, Lemma 16], this implies that any limit point  $\mathbf{z}_\infty$  of  $(\mathbf{z}_k)_{k \geq k_0}$  satisfies  $\mathbf{z}_\infty \in \overline{D}$ . This is a contradiction with the fact that  $\mathbf{z}_k - \xi \mathbf{j} \notin D$  for every  $k \geq k_0$ .  $\square$

The following lemma shows that if  $[\alpha]^w$  is  $(\gamma, \epsilon)$ -good, then the evaluation of  $Y^u$  in a point close to  $\mathbf{z}_{n,\alpha}$  is close to  $\frac{1}{2}$  (which was shown in [CP15] to be the maximum of  $Y^u$  on  $D$ ).

**Lemma 9.** *For every  $\rho$ , every  $\epsilon$  and every  $\ell$  such that  $\ell < 1/\epsilon$ , if  $\gamma \leq \gamma_0(\rho, \ell)$ ,  $\xi \leq \xi_0(\rho, \epsilon, \ell)$ ,  $k_* \geq k_0(\epsilon, \xi)$  and  $n$  is large enough, then for every box  $[\alpha]^w$  which is  $(\gamma, \epsilon)$ -good the following holds for  $\hat{\mathbf{z}} := \mathbf{z}_{n,\alpha} - \xi \mathbf{j}$ : we have  $\hat{\mathbf{z}} \in D$ ,*

$$Y^u(\hat{\mathbf{z}}) > \frac{1}{2} - \rho,$$

and for every  $U \in \mathcal{U}_{\leq \ell}$ , we have

$$|\omega(U, \hat{\mathbf{z}}) - \hat{z}^U| \leq \rho.$$

*Proof.* Let  $\gamma_0 := \frac{\rho}{4\ell!}$  and  $\xi_0 := \frac{\rho}{2|\mathcal{U}_\epsilon|\ell!}$ . Consider  $\alpha \in \mathcal{E}_{k_*}$  such that the box  $[\alpha]^w$  is  $(\gamma, \epsilon)$ -good. Using the properties i) and ii) of good boxes, and (20), we have

$$\begin{aligned} \tilde{Y}_{\mathcal{U}_\epsilon}^u(\mathbf{z}_{n,\alpha}) &= \sum_{U \in \mathcal{U}_\epsilon} \frac{z_{n,\alpha}^U}{\text{Aut}_u(U)} = \frac{1}{|\mathcal{A}_{n,[\alpha]}^w|} \sum_{U \in \mathcal{U}_\epsilon} |\mathcal{B}_{n,[\alpha]}^U| \left(1 - \frac{|U|}{n}\right) \\ &\geq \frac{1}{|\mathcal{A}_{n,[\alpha]}^w|} |\mathcal{B}_{n,[\alpha]}^w| (1 - \gamma) \left(1 - \frac{|U|}{n}\right) \\ &\geq \left(\frac{1}{2} - \gamma\right) (1 - \gamma) \left(1 - \frac{1}{\epsilon n}\right) \\ &\geq \frac{1}{2} - 2\gamma, \end{aligned}$$

provided that  $n$  is large enough. Now, since  $\tilde{Y}_{\mathcal{U}_\epsilon}^u(\hat{\mathbf{z}})$  is a finite sum, we have

$$\tilde{Y}_{\mathcal{U}_\epsilon}^u(\hat{\mathbf{z}}) \geq \tilde{Y}_{\mathcal{U}_\epsilon}^u(\mathbf{z}_{n,\alpha}) - \xi |\mathcal{U}_\epsilon|.$$

Together with the previous inequality and the choice of  $\gamma_0$  and  $\xi_0$ , this implies

$$\tilde{Y}_{\mathcal{U}_\epsilon}^u(\hat{\mathbf{z}}) \geq \frac{1}{2} - (\xi|\mathcal{U}_\epsilon| + 2\gamma) \geq \frac{1}{2} - \frac{\rho}{\ell!} . \quad (21)$$

By definition of maximum weight, for every  $U \in \mathcal{U}_\epsilon$  we have  $\omega(U, \mathbf{z}) \geq z^U$ , which directly implies  $Y_{\mathcal{U}_\epsilon}^u(\mathbf{z}) \geq \tilde{Y}_{\mathcal{U}_\epsilon}^u(\mathbf{z})$ . We thus conclude the first part of the lemma,

$$Y^u(\hat{\mathbf{z}}) \geq Y_{\mathcal{U}_\epsilon}^u(\hat{\mathbf{z}}) \geq \tilde{Y}_{\mathcal{U}_\epsilon}^u(\hat{\mathbf{z}}) \geq \frac{1}{2} - \frac{\rho}{\ell!} > \frac{1}{2} - \rho$$

Observe that this is true even if  $\hat{\mathbf{z}} \notin D$ , since then the LHS is infinite.

By Lemma 8, we can choose  $k_0 = k_0(\epsilon, \xi)$  such that if  $k_* \geq k_0$  and  $n$  is large enough, we have  $\hat{\mathbf{z}} \in D$ . The choice of  $k_*$  and  $n$  is suitable for all vectors in  $\mathcal{E}_{k_*}$ . Then, Lemma 14 in [CP15] implies that  $Y_{\mathcal{U}_\epsilon}^u(\hat{\mathbf{z}}) \leq Y^u(\hat{\mathbf{z}}) \leq \frac{1}{2}$ . Together with (21), for every  $U \in \mathcal{U}_\epsilon$  we have

$$\frac{\rho}{\ell!} \geq |Y_{\mathcal{U}_\epsilon}^u(\hat{\mathbf{z}}) - \tilde{Y}_{\mathcal{U}_\epsilon}^u(\hat{\mathbf{z}})| = \left| \sum_{U' \in \mathcal{U}_\epsilon} \frac{\omega(U', \hat{\mathbf{z}}) - \hat{z}^{U'}}{\text{Aut}_u(U')} \right| \geq \frac{|\omega(U, \hat{\mathbf{z}}) - \hat{z}^U|}{\text{Aut}_u(U)} ,$$

where the last inequality follows since  $\omega(U', \hat{\mathbf{z}}) \geq \hat{z}^{U'}$  for each tree  $U' \in \mathcal{U}_\epsilon$ . Since  $\text{Aut}_u(U) \leq \ell!$ , it follows that

$$|\omega(U, \hat{\mathbf{z}}) - \hat{z}^U| \leq \rho . \quad \square$$

The next lemma states that if  $\mathbf{z}$  belongs to  $D$  and  $Y^u(\mathbf{z})$  is close to  $\frac{1}{2}$ , then  $\omega(T, \mathbf{z})$  is close to  $e^{-|T|}$  for every  $T$  with bounded size.

**Lemma 10.** *For every  $\nu$  and every  $\ell$ , if  $\rho \leq \rho_0(\nu, \ell)$ , then for every  $\epsilon$ , every  $\mathbf{z} \in D$  that satisfies  $Y^u(\mathbf{z}) > \frac{1}{2} - \rho$ , and every  $T \in \mathcal{T}_{\leq \ell}$ , we have*

$$|\omega(T, \mathbf{z}) - e^{-|T|}| < \nu . \quad (22)$$

*Proof.* Let  $Y^e(\mathbf{z})$  be the partition function of trees rooted at an edge, where each tree is weighted by its maximal weight. As noted in [CP15], a classical trick known as the dissymmetry theorem [BLL98] implies that

$$Y^e(\mathbf{z}) = Y(\mathbf{z}) - Y^u(\mathbf{z}) .$$

Together with the hypothesis of the lemma and the fact that  $y - 1/2 \leq y^2/2$  for all  $y \in \mathbb{R}$ , this implies:

$$Y^e(\mathbf{z}) = Y(\mathbf{z}) - Y^u(\mathbf{z}) \leq Y(\mathbf{z}) - 1/2 + \rho \leq \frac{1}{2}(Y(\mathbf{z}))^2 + \rho ,$$

For every pair of vertex rooted trees  $T_1, T_2 \in \mathcal{T}$ , let  $f(T_1, T_2)$  be the edge-rooted tree obtained by adding an edge (the root) connecting the roots of  $T_1$  and  $T_2$ . We have the following supermultiplicativity property:

$$\omega(f(T_1, T_2), \mathbf{z}) - \omega(T_1, \mathbf{z})\omega(T_2, \mathbf{z}) \geq 0 .$$



Also observe that the number of automorphisms of  $f(T_1, T_2)$  that fix the rooted edge (as an ordered edge!), is precisely  $\text{Aut}_r(T_1)\text{Aut}_r(T_2)$ . Thus, for any pair  $R_1, R_2 \in \mathcal{T}$ , we have

$$\begin{aligned} \rho &\geq Y^e(\mathbf{z}) - \frac{1}{2}(Y(\mathbf{z}))^2 = \sum_{T_1, T_2 \in \mathcal{T}} \frac{\omega(f(T_1, T_2), \mathbf{z}) - \omega(T_1, \mathbf{z})\omega(T_2, \mathbf{z})}{\text{Aut}_r(T_1)\text{Aut}_r(T_2)} \\ &\geq \frac{\omega(f(R_1, R_2), \mathbf{z}) - \omega(R_1, \mathbf{z})\omega(R_2, \mathbf{z})}{|R_1|! |R_2|!}. \end{aligned} \quad (23)$$

Let  $\bullet$  be the tree composed by a single vertex and define  $x = x(\mathbf{z}) := \omega(\bullet, \mathbf{z}) = z^\bullet \in \mathbb{R}_+$ . Observe that since  $\mathbf{z} \in D$ , we have  $x \leq 1$  (otherwise  $Y(\mathbf{z}) = \infty$  since  $\omega(T, \mathbf{z}) \geq x^{|T|}$ ). Using (23) with  $R_2 = \bullet$ , for every  $T \in \mathcal{T}$ :

$$\omega(f(T, \bullet), \mathbf{z}) \leq x \cdot \omega(T, \mathbf{z}) + \rho \cdot |T|!,$$

and induction on  $|T|$  implies that for every  $T \in \mathcal{T}$  we have

$$x^{|T|} \leq \omega(T, \mathbf{z}) \leq x^{|T|} + |T|! \rho \leq \left( x + (\rho |T|!)^{\frac{1}{|T|}} \right)^{|T|}.$$

Note that if  $|T| \leq \ell$ , then  $(\rho |T|!)^{\frac{1}{|T|}} \leq c(\ell) \rho^{\frac{1}{\ell}}$ , for some  $c(\ell) > 0$ . Consider  $\mathbf{x} = (x^{|U|})_{U \in \mathcal{U}_\ell}$  and  $\mathbf{x}_\rho = ((x + c(\ell) \rho^{\frac{1}{\ell}})^{|U|})_{U \in \mathcal{U}_\ell}$ . By the definition of  $\mathbf{x}$ , note that  $\omega(T, \mathbf{x}) = x^{|T|}$ , therefore  $\omega(T, \mathbf{x}) \leq \omega(T, \mathbf{z})$  and since  $\mathbf{z} \in D$ , by Lemma 14 in [CP15] we have

$$Y^u(\mathbf{x}) \leq Y^u(\mathbf{z}) \leq \frac{1}{2}.$$

This implies  $x \leq e^{-1}$  (otherwise  $Y^u(\mathbf{x})$  would not converge). Similarly  $\omega(T, \mathbf{x}_\rho) = (x + c(\ell) \rho^{\frac{1}{\ell}})^{|T|}$ , and using the hypothesis of the lemma we have

$$\frac{1}{2} - \rho \leq Y^u(\mathbf{z}) \leq Y^u(\mathbf{x}_\rho).$$

By Equation (12) in Lemma 6, this implies that  $x + c(\ell) \rho^{\frac{1}{\ell}} \geq e^{-1} - O(\sqrt{c(\ell) \rho^{1/\ell}})$ . Given  $\nu$  and  $\ell$ , we can now set  $\rho_0(\nu, \ell)$  small enough such that for  $\rho \leq \rho_0(\nu, \ell)$  one has  $x > e^{-1} (1 - y)$ , with  $y = \min\{\frac{\nu e^\ell}{\ell}, 1\}$ , and  $\rho \leq \frac{\nu}{\ell}$ . We then have, for every  $T \in \mathcal{T}_{\leq \ell}$ ,

$$e^{-|T|} - \nu \leq e^{-|T|} (1 - y |T|) \leq e^{-|T|} (1 - y)^{|T|} < x^{|T|} \leq \omega(T, \mathbf{z}) \leq x^{|T|} + \rho \cdot |T|! \leq e^{-|T|} + \nu.$$

where we used that  $(1 - y)^\ell$  is convex for  $y \in [0, 1]$ .  $\square$

Finally, we can prove estimates for the ratios between  $|\mathcal{B}_{n, [\alpha]^w}^U|$  and  $|\mathcal{A}_{n, [\alpha]_q^w}|$  for good boxes  $[\alpha]^w$  and unrooted trees  $U$  with bounded size.

**Proposition 11.** *For every  $\vartheta$ , every  $\epsilon$  and every  $\ell$  such that  $\ell < 1/\epsilon$ , if  $\gamma \leq \gamma_0(\vartheta, \ell)$ ,  $k_* \geq k_0(\vartheta, \epsilon, \ell)$  and  $n$  is large enough, then for every box  $[\alpha]^w$  which is  $(\gamma, \epsilon)$ -good and every  $U \in \mathcal{U}_{\leq \ell}$*

$$\left| \frac{|\mathcal{B}_{n, [\alpha]^w}^U|}{|\mathcal{A}_{n, [\alpha]_q^w}|} - \frac{e^{-|U|}}{\text{Aut}_u(U)} \right| < \vartheta.$$

*Proof.* Let us first fix the constants that we will need in the proof. For  $\nu := \vartheta/4$ , we let  $\rho_0 = \rho_0(\nu, \ell)$  be the value obtained from Lemma 10. For  $\rho := \min\{\rho_0, \nu\}$ , we let  $\gamma_0 = \gamma_0(\rho, \ell)$ ,  $\xi_0 = \xi_0(\rho, \epsilon, \ell)$  be the values obtained from Lemma 9. For  $\xi := \min\{\xi_0, \nu\}$ , we let  $k_0 = k_0(\epsilon, \xi)(= k_0(\vartheta, \epsilon, \ell))$  be the value obtained from Lemma 9. Now fix  $k_* \geq k_0$  and consider  $n$  large enough. Note that once  $k_*$  and  $n$  are chosen, the space  $\mathcal{E}_{k_*}$  is well-determined.

Let  $\hat{\mathbf{z}} = \mathbf{z}_{n,\alpha} - \xi \mathbf{j}$  as before. For a given  $U \in \mathcal{U}_{\leq \ell}$ , we observe

$$|z_{n,\alpha}^U - \hat{z}^U| \leq \xi \leq \vartheta/4.$$

By Lemma 9, if  $[\alpha]^w$  is  $(\gamma, \epsilon)$ -good, we have

$$|\hat{z}^U - \omega(U, \hat{\mathbf{z}})| \leq \rho \leq \vartheta/4.$$

The same lemma also implies that  $\hat{\mathbf{z}} \in D$  and that  $Y^u(\hat{\mathbf{z}}) > \frac{1}{2} - \rho$ . Thus,  $\hat{\mathbf{z}}$  satisfies the hypothesis of Lemma 10, which implies

$$|\omega(U, \hat{\mathbf{z}}) - e^{-|U|}| < \nu = \vartheta/4.$$

Using the previous three inequalities and (20), we conclude

$$\left| \frac{|\mathcal{B}_{n,[\alpha]^w}^U|}{|\mathcal{A}_{n,[\alpha]^w}|} - \frac{e^{-|U|}}{\text{Aut}_u(U)} \right| = \frac{|z_{n,\alpha}^U \left(1 - \frac{|U|}{n}\right)^{-1} - e^{-|U|}|}{\text{Aut}_u(U)} < \vartheta,$$

provided that  $n$  is large enough. In the last inequality we used that  $z_{n,\alpha}^U \leq 1$  (this can be obtained using a similar argument as the one used to obtain (7)).  $\square$

**Proposition 12.** *For every  $\vartheta$ , every  $\epsilon$  and every  $\ell$  such  $\ell < 1/\epsilon$ , if  $\gamma \leq \gamma_0(\vartheta, \ell)$ ,  $k_* \geq k_0(\vartheta, \epsilon, \ell)$  and  $n$  is large enough, then for every box  $[\alpha]^w$  which is  $(\gamma, \epsilon)$ -good and every  $T \in \mathcal{T}_{\leq \ell}$*

$$\left| \frac{\alpha(T)}{n} - \frac{e^{-|T|}}{\text{Aut}_r(T)} \right| < \vartheta.$$

*Proof.* Again, let us start by fixing the constants that we will need in the proof. For  $\nu := \frac{\vartheta}{4|\mathcal{T}_{\leq \ell}|}$ , we let  $\rho_0 = \rho_0(\nu, \ell)$  be the value obtained from Lemma 10. For  $\rho \leq \rho_0$ , we let  $\gamma_0 = \gamma_0(\rho, \ell)(= \gamma_0(\vartheta, \ell))$ ,  $\xi_0 = \xi_0(\rho, \epsilon, \ell)$  be the values obtained from Lemma 9.

Observe that, if we fix  $T \in \mathcal{T}_{\leq \ell}$ , the function  $\omega(T, \mathbf{z})$  is a *piecewise* polynomial in the set of variables  $\{z^U : U \in \mathcal{U}_\epsilon\}$  that it is continuous at every point of  $(\mathbb{R}_+)^{\mathcal{U}_\epsilon}$ . Since  $D$  is bounded, there exists a  $\xi_1$  such that for every  $\xi \leq \xi_1$  and every  $\mathbf{z}$  at distance at most 1 from  $D$  (in the  $\ell_\infty$  norm), we have

$$|\omega(T, \mathbf{z}) - \omega(T, \mathbf{z} - \xi \mathbf{j})| < \frac{\vartheta}{4|\mathcal{T}_{\leq \ell}|}.$$

For  $\xi := \min\{\xi_0, \xi_1\}$ , we let  $k_0 = k_0(\epsilon, \xi)(= k_0(\vartheta, \epsilon, \ell))$  be the value obtained from Lemma 9. Fix  $k_* \geq k_0$  and consider  $n$  large enough.

By Lemma 9, if  $[\alpha]^w$  is  $(\gamma, \epsilon)$ -good and we write  $\hat{\mathbf{z}} := \mathbf{z}_{n,\alpha} - \xi \mathbf{j}$ , we have  $\hat{\mathbf{z}} \in D$  and  $Y^u(\hat{\mathbf{z}}) > \frac{1}{2} - \rho$ . Thus,  $\hat{\mathbf{z}}$  satisfies the hypothesis of Lemma 10 and we have

$$|\omega(T, \hat{\mathbf{z}}) - e^{-|T|}| < \nu = \frac{\vartheta}{4|\mathcal{T}_{\leq \ell}|}.$$

Using the previous inequalities, we obtain

$$|\omega(T, \mathbf{z}_{n,\alpha}) - e^{-|T|}| \leq |\omega(T, \mathbf{z}_{n,\alpha}) - \omega(T, \hat{\mathbf{z}})| + |\omega(T, \hat{\mathbf{z}}) - e^{-|T|}| < \frac{\vartheta}{2|\mathcal{T}_{\leq \ell}|}.$$

By Lemma 11 in [CP15], there exists a constant  $C$  that does not depend on  $n$  such that

$$\frac{\alpha(T)}{n} \geq \frac{\omega(T, \mathbf{z}_{n,\alpha})}{\text{Aut}_r(T)} - \frac{C}{n} \geq \frac{e^{-|T|}}{\text{Aut}_r(T)} - \frac{2\vartheta}{3|\mathcal{T}_{\leq \ell}|}, \quad (24)$$

where the last inequality holds provided  $n$  is large enough. This proves one side of the inequality in the statement.

By Lemma 6, if we let  $t$  be large enough with respect to  $\vartheta$ , we have that

$$\sum_{T \in \mathcal{T}_{\leq t}} \frac{e^{-|T|}}{\text{Aut}_r(T)} > 1 - \frac{\vartheta}{3}. \quad (25)$$

We can assume that  $\ell \geq t$ , up to increasing the value of  $k_*$  and  $n$ .

For the sake of contradiction, suppose that there exists  $T_0 \in \mathcal{T}_{\leq \ell}$  such that  $\frac{\alpha(T_0)}{n} > \frac{e^{-|T_0|}}{\text{Aut}_r(T_0)} + \vartheta$ . Then, using (24), (25) and the properties of  $T_0$ , we get

$$1 \geq \sum_{T \in \mathcal{T}_{\leq \ell}} \frac{\alpha(T)}{n} \geq \sum_{T \in \mathcal{T}_{\leq \ell}} \frac{e^{-|T|}}{\text{Aut}_r(T)} - \frac{2\vartheta}{3} + \vartheta > 1,$$

thus obtaining a contradiction and concluding the proof of the lemma.  $\square$

### 3.3 Proof of Theorem 1 for classes of forests: the case of 1 or 2 connected components

For every  $\delta$  and every  $\ell$ , consider the set of vectors in  $\mathcal{E}_\ell$  that are  $\delta$ -close from the distribution  $a_\infty$  (recall that for  $T \in \mathcal{T}$ ,  $a_\infty(T) = \frac{e^{-|T|}}{\text{Aut}_r(T)}$ ); that is,

$$\Xi(\delta, \ell) = \left\{ \beta \in \mathcal{E}_\ell : \left| \frac{\beta(T)}{n} - a_\infty(T) \right| < \delta, \text{ for every } T \in \mathcal{T}_{\leq \ell} \right\}. \quad (26)$$

In what follows, for every set of graphs  $\mathcal{S}_n$ , every  $\ell \geq 1$  and every  $\beta \in \mathcal{E}_\ell$ , we let  $\mathcal{S}_{n,\beta}$  be the set of graphs  $G$  in  $\mathcal{S}_n$  such that  $\alpha^G(T) = \beta(T)$  for all  $T \in \mathcal{T}_{\leq \ell}$ .

**Proposition 13.** *For every  $\theta_1$  and every  $U \in \mathcal{U}$ , there exists a  $\zeta$  such that for every  $\zeta$ -tight class  $\mathcal{G}$  of forests and every large enough  $n$ , one has*

$$\left| \frac{|\mathcal{B}_n^U|}{|\mathcal{G}_n|} - e^{-1/2} \frac{e^{-|U|}}{\text{Aut}_u(U)} \right| < \theta_1.$$

Moreover, for every  $\theta_1$ , every  $\delta$ , every  $\ell$  and every  $U \in \mathcal{U}$ , there exists a  $\zeta$  such that for every  $\zeta$ -tight class  $\mathcal{G}$  of forests and every large enough  $n$ , one has

$$\left| \frac{\sum_{\beta \in \Xi(\delta, \ell)} |\mathcal{B}_{n,\beta}^U|}{|\mathcal{B}_n^U|} - 1 \right| < \theta_1.$$

*Proof.* We start by fixing the constants needed in the proof. For  $\vartheta := \theta_1/8$  and  $\ell = |U|$ , we let  $\gamma_0 = \gamma_0(\vartheta, \ell)$  be the constant obtained from Proposition 11. Fix  $\gamma \leq \gamma_0$ . For  $\eta := \frac{\theta_1}{4}$ , we let  $\epsilon_0 = \epsilon_0(\gamma, \eta)$  be the constant obtained from Lemma 7. For  $\epsilon := \min\{\epsilon_0, 1/\ell, \theta_1/8\}$ , we let  $k_0(\vartheta, \epsilon, \ell)$  be the maximum of the constants obtained from Lemma 7 and Proposition 11. Fix  $k_* \geq k_0$ . Let  $\zeta$  be the minimum between the constant obtained from Lemma 7 and  $\theta_1/8$ . Let  $n$  be large enough with respect to all the previous parameters.

Now that  $\epsilon$  and  $k_*$  are fixed, we consider as before the family  $\mathcal{U}_\epsilon \subset \mathcal{U}$  of unrooted trees of order at most  $\lceil \epsilon^{-1} \rceil$  and the family  $\mathcal{T}_* \subset \mathcal{T}$  of all rooted trees of order at most  $k_*$ . We also let  $w$  and  $K$ , and the collection of boxes  $\{[\beta_i]^w, 1 \leq i \leq K\}$  be defined (relatively to the values of  $\epsilon$  and  $k_*$ ) as in Section 3.1. We recall that these boxes satisfy (14), and using (13) we note that  $\sum_{i=1}^K |\mathcal{A}_{n, [\beta_i]^w}| = |\mathcal{A}_n|$ .

We can write,

$$\begin{aligned} \left| \frac{|\mathcal{B}_n^U|}{|\mathcal{G}_n|} - e^{-1/2} \frac{e^{-|U|}}{\text{Aut}_u(U)} \right| &\leq \left| \sum_{i=1}^K \frac{|\mathcal{B}_{n, [\beta_i]^w}^U|}{|\mathcal{G}_n|} - e^{-1/2} \frac{e^{-|U|}}{\text{Aut}_u(U)} \right| + \epsilon \\ &\leq \left| \sum_{i \notin \text{Good}_{\gamma, \epsilon}} \frac{|\mathcal{B}_{n, [\beta_i]^w}^U|}{|\mathcal{G}_n|} + \frac{1}{|\mathcal{G}_n|} \sum_{i \in \text{Good}_{\gamma, \epsilon}} |\mathcal{B}_{n, [\beta_i]^w}^U| - e^{-1/2} \frac{e^{-|U|}}{\text{Aut}_u(U)} \right| + \frac{\theta_1}{8}. \end{aligned}$$

By Proposition 11, for every  $i \in \text{Good}_{\gamma, \epsilon}$  and every  $U \in \mathcal{U}_{\leq \ell}$ , we have

$$\left| |\mathcal{B}_{n, [\beta_i]^w}^U| - \frac{e^{-|U|}}{\text{Aut}_u(U)} \cdot |\mathcal{A}_{n, [\beta_i]^w}| \right| \leq \frac{\theta_1}{8}.$$

By Lemma 7, we have

$$\sum_{i \notin \text{Good}_{\gamma, \epsilon}} \frac{|\mathcal{B}_{n, [\beta_i]^w}^U|}{|\mathcal{G}_n|} \leq \sum_{i \notin \text{Good}_{\gamma, \epsilon}} \frac{|\mathcal{B}_{n, [\beta_i]^w}|}{|\mathcal{B}_n|} \leq \eta = \frac{\theta_1}{4}.$$

Let  $M$  be the number of boxes  $[\beta_i]$  that are non-empty. Clearly,  $M \leq |\mathcal{G}_n|$ . Therefore,

$$\begin{aligned} &\left| \frac{|\mathcal{B}_n^U|}{|\mathcal{G}_n|} - e^{-1/2} \frac{e^{-|U|}}{\text{Aut}_u(U)} \right| \\ &\leq \frac{\theta_1}{4} + \frac{\theta_1 M}{8|\mathcal{G}_n|} + \left| \frac{e^{-|U|}}{\text{Aut}_u(U)|\mathcal{G}_n|} \left( \sum_{i \in \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]^w}| \right) - e^{-1/2} \frac{e^{-|U|}}{\text{Aut}_u(U)} \right| + \frac{\theta_1}{8} \\ &\leq \frac{\theta_1}{2} + \left| \frac{|\mathcal{A}_n|}{|\mathcal{G}_n|} \left( \frac{\sum_{i \in \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]^w}|}{|\mathcal{A}_n|} \right) - e^{-1/2} \right| \frac{e^{-|U|}}{\text{Aut}_u(U)} \end{aligned}$$

Again, by Lemma 7 and using that  $\sum_{i=1}^K |\mathcal{A}_{n, [\beta_i]^w}| = |\mathcal{A}_n|$ , we have

$$\left| \frac{\sum_{i \in \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]^w}|}{|\mathcal{A}_n|} - 1 \right| = \frac{\sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{A}_{n, [\beta_i]^w}|}{|\mathcal{A}_n|} \leq \eta = \frac{\theta_1}{4}.$$

Since  $\mathcal{G}$  is a  $\zeta$ -tight bridge-addable class, by definition, using Theorem A and provided that  $n$  is large enough, we obtain

$$(1 - \zeta)e^{-1/2} \leq \frac{|\mathcal{A}_n|}{|\mathcal{G}_n|} \leq (1 + \zeta)e^{-1/2}.$$

Since  $\zeta \leq \theta_1/8$ , we obtain

$$\begin{aligned} \left| \frac{|\mathcal{B}_n^U|}{|\mathcal{G}_n|} - e^{-1/2} \frac{e^{-|U|}}{\text{Aut}_u(U)} \right| &\leq \frac{\theta_1}{2} + \left( \left(1 + \frac{\theta_1}{8}\right) \left(1 + \frac{\theta_1}{4}\right) - 1 \right) e^{-1/2} \frac{e^{-|U|}}{\text{Aut}_u(U)} \\ &\leq \frac{\theta_1}{2} + \frac{\theta_1}{2} \cdot e^{-1/2} \frac{e^{-|U|}}{\text{Aut}_u(U)} \leq \theta_1. \end{aligned}$$

This concludes the proof of the first part of the proposition.

For the second part, let us proceed by contradiction. Suppose that there exist  $\theta, \delta, \ell$  and  $U \in \mathcal{U}$ , such that for every  $\zeta$  there exists a  $\zeta$ -tight class  $\mathcal{G}$  and a large enough  $n$  with

$$\left| \frac{\sum_{\beta \in \Xi(\delta, \ell)} |\mathcal{B}_{n, \beta}^U|}{|\mathcal{B}_n^U|} - 1 \right| > \theta.$$

or equivalently,

$$\frac{\sum_{\beta \notin \Xi(\delta, \ell)} |\mathcal{B}_{n, \beta}^U|}{|\mathcal{B}_n^U|} > \theta. \quad (27)$$

Note that by the first part of the proposition with  $\theta_1$  small enough, we have that  $\frac{|\mathcal{B}_n^U|}{|\mathcal{G}_n|}$  is arbitrarily close to  $e^{-1/2} \frac{e^{-|U|}}{\text{Aut}_u(U)}$ , for  $\zeta$  small and  $n$  large enough. Thus, there exists a uniform constant  $c(U) > 0$  such that  $\frac{|\mathcal{B}_n^U|}{|\mathcal{B}_n|} \geq c(U)$ , and (27) is well-defined.

Let  $\eta = \theta c(U)$  and let  $\vartheta = \delta/2$ . As in the first part of the proposition, we can choose  $\gamma, \epsilon, k_*, \zeta$  and  $n$ , such that Lemma 7 and Proposition 12 can be applied. We skip the details of this setting. We will again consider the set of boxes  $\{[\beta_i] : 1 \leq i \leq K\}$  of  $\mathcal{E}_{k_*}$  fixed in Section 3.1. For every  $\alpha \in \mathcal{E}_{k_*}$ , we consider its canonical projection  $\pi(\alpha)$  onto  $\mathcal{E}_\ell$  obtained by selecting the first  $|\mathcal{E}_\ell|$  coordinates of  $\alpha$ .

**Claim.** Let  $\alpha \in [\beta_i]^w$ , for some  $i \in \text{Good}(\gamma, \epsilon)$ . Then  $\pi(\alpha) \in \Xi(\delta, \ell)$ .

*Proof of the Claim.* By Proposition 12 and since  $[\beta_i]$  is  $(\gamma, \epsilon)$ -good, for every  $T \in \mathcal{T}_{\leq \ell}$  we have

$$\left| \frac{\beta_i(T)}{n} - \frac{e^{-|T|}}{\text{Aut}_r(T)} \right| < \vartheta.$$

Since  $\alpha \in [\beta_i]^w$ , for every  $T \in \mathcal{T}_{\leq k_*}$ , we have  $|\beta_i(T) - \alpha(T)| \leq w$ . The choice of  $w$  does not depend on  $n$ , and thus,  $\left| \frac{\beta_i(T)}{n} - \frac{\alpha(T)}{n} \right| \leq \frac{\delta}{3}$ , if  $n$  large enough. Since  $\ell \leq k_*$ , for every  $T \in \mathcal{T}_{\leq \ell}$  we have

$$\left| \frac{\alpha(T)}{n} - \frac{e^{-|T|}}{\text{Aut}_r(T)} \right| < \vartheta + \frac{\delta}{3} < \delta.$$

We conclude that  $\pi(\alpha) \in \Xi(\delta, \ell)$ , which proves the claim.  $\square$

As a direct corollary of the claim, we get

$$\frac{\sum_{\beta \notin \Xi(\delta, \ell)} |\mathcal{B}_{n, \beta}^U|}{|\mathcal{B}_n^U|} \leq \frac{\sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{B}_{n, [\beta_i]^w}^U|}{|\mathcal{B}_n^U|}.$$

By Lemma 7, it follows that

$$\frac{\sum_{\beta \notin \Xi(\delta, \ell)} |\mathcal{B}_{n, \beta}^U|}{|\mathcal{B}_n^U|} \leq \frac{|\mathcal{B}_n|}{|\mathcal{B}_n^U|} \cdot \frac{\sum_{i \notin \text{Good}_{\gamma, \epsilon}} |\mathcal{B}_{n, [\beta_i]^w}^U|}{|\mathcal{B}_n|} \leq \frac{|\mathcal{B}_n|}{|\mathcal{B}_n^U|} \cdot \eta \leq \theta,$$

where we have used  $|\mathcal{B}_{n, [\beta_i]^w}^U| \leq |\mathcal{B}_{n, [\beta_i]^w}|$ , giving a contradiction with (27). □

### 3.4 Proof of Theorem 1 for classes of forests

We now prove the main result of this section, Theorem 14, that is equivalent to our main theorem for bridge-addable classes of forests.

Here we need to introduce the notion of *removable edges*. This notion will also be crucial in the next section in order to transfer the obtained result from classes of forests to general graph classes. We say that an edge  $e$  in a graph  $G \in \mathcal{G}$  is *removable* if the graph  $G' = G \setminus e$  is in  $\mathcal{G}$ . For a subclass  $\mathcal{H} \subseteq \mathcal{G}$  and a rooted tree  $T \in \mathcal{T}$ , we define  $p(\mathcal{H}, T)$  to be the probability that given a uniform random graph  $H \in \mathcal{H}$ , and a uniform random pendant copy of  $T$  in  $H$ , the graph  $H'$  obtained by deleting the edge that connects the pendant copy of  $T$  to the rest of the graph belongs to  $\mathcal{G}$  (and not only to  $\mathcal{H}$ ). In other words,  $p(\mathcal{H}, T)$  is the average over all graphs in  $\mathcal{H}$  of the proportion of pendant copies of  $T$  that are attached using a removable edge. This notion is inspired by bridge-alterable classes, for which  $p(\mathcal{H}, T) = 1$ , for every  $\mathcal{H} \subseteq \mathcal{G}$  and every  $T \in \mathcal{T}$  [ABMR12, KP13]. We do a slight abuse of notation by writing  $p(G, T)$  for  $p(\{G\}, T)$ , for each  $G \in \mathcal{G}$ . Also, in the cases where  $p(G, T)$  is not well-defined (that is, if  $G$  has no pendant copy of  $T$ ), we interpret the probability as 1.

Recall the definition of  $\Xi(\delta, \ell)$  given in (26), and recall from Section 2.2 that we use the notation  $\{U_1, U_2, \dots, U_k\}$  to denote the forest formed by a *multiset* of  $k$  unrooted trees.

**Theorem 14.** *For every  $k \geq 1$ , every  $\theta_k$  and every  $U_1, \dots, U_k \in \mathcal{U}$ , there exists a  $\zeta$  such that for every  $\zeta$ -tight class  $\mathcal{G}$  of forests and every large enough  $n$ , one has*

$$\left| \frac{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n|} - e^{-1/2} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_1, \dots, U_k)} \right| < \theta_k. \quad (28)$$

Moreover, for every  $k$ , every  $\ell$ , every  $\theta_k$ , every  $\delta$  and every  $U_1, \dots, U_k \in \mathcal{U}$ , there exists a  $\zeta$  such that for every  $\zeta$ -tight class  $\mathcal{G}$  of forests and every large enough  $n$ , one has

$$\left| \frac{\sum_{\beta \in \Xi(\delta, \ell)} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|} - 1 \right| < \theta_k. \quad (29)$$

**Note:** In the following proof and in the rest of the paper, we will often partition the set  $\mathcal{G}_n$  of graphs in  $\mathcal{G}$  having  $n$  vertices into smaller graph classes. If  $\mathcal{H} \subset \mathcal{G}_n$  is a part in this partition, and  $\zeta > 0$  is a real number, we will say that  $\mathcal{H}$  is  $\zeta$ -tight if

$$\Pr(H \text{ is connected}) \leq (1 + \zeta)e^{-1/2},$$

where  $H$  is uniform in  $\mathcal{H}$ . In other words, we slightly adapt the notion of  $\zeta$ -tightness for classes of graphs which from the context have a fixed number of vertices.

*Proof of Theorem 14.* We will start by showing the first part of the theorem by induction on  $k$ , and then prove the second part, again by induction on  $k$  and using the result obtained in the first part. The proof of the first part will also require the proof of an intermediate claim.

**Proof of the first part.** Fix  $U_1, \dots, U_k \in \mathcal{U}$  and let  $u$  be an upper bound on their sizes.

We are going to prove the first statement by induction. Observe that Proposition 13 proves the case  $k = 1$ . Assume that the statement is true for  $k - 1$  and let us show it for  $k$ .

We consider the following total order on the subsets of  $[n]$ ; for every  $V_1, V_2 \subseteq V$  we have  $V_1 < V_2$  if  $|V_1| < |V_2|$  or  $|V_1| = |V_2|$  and the elements of  $V_1$  are smaller in the lexicographical order, than the ones in  $V_2$ .

Let  $m(U_1, \dots, U_k)$  be the number of graphs isomorphic to  $U_k$  among  $U_1, \dots, U_k$ . Observe that

$$\text{Aut}_u(U_1, \dots, U_k) = m(U_1, \dots, U_k) \text{Aut}_u(U_k) \text{Aut}_u(U_1, \dots, U_{k-1}). \quad (30)$$

For every subset of vertices  $W \subset V$ , we use  $G[W]$  to denote the graph induced by  $W$  in  $G$ . For every unlabeled graph  $U$ , the notation  $G[W] \equiv U$ , not only denotes graph isomorphism, but also that  $W$  induces a maximal connected component in  $G$ .

Given disjoint sets  $V_1, \dots, V_{k-1} \subset [n]$ , consider the graph class

$$\mathcal{H}(V_1, \dots, V_{k-1}) = \{G[V \setminus \cup_{i=1}^{k-1} V_i] : G \in \mathcal{G}_n, G[V_1] \equiv U_1, \dots, G[V_{k-1}] \equiv U_{k-1}\}.$$

In what follows, given  $U_1, \dots, U_{k-1}$ , we will consider all the possible tuples of subsets  $V_1, \dots, V_{k-1}$  underlying the small induced components of our graph. In order to avoid multiplicity problems and avoid considering the same tuple several times, we need to carefully define the set of such tuples in the case where several  $U_i$ 's are isomorphic to each other. To this end, we consider the set of  $(k - 1)$ -tuples of disjoint subsets defined as follows,

$$\mathcal{V} = \{(V_1, \dots, V_{k-1}), V_i \subset [n] \text{ disjoint; if } U_i \equiv U_j \text{ then } V_i < V_j\}. \quad (31)$$

We then write  $\mathcal{H} = \cup_{(V_1, \dots, V_{k-1}) \in \mathcal{V}} \mathcal{H}(V_1, \dots, V_{k-1})$ .

Since  $\mathcal{G}$  is a bridge-addable class on the set of vertices  $V$ , then  $\mathcal{H}(V_1, \dots, V_{k-1})$  (for every  $(V_1, \dots, V_{k-1}) \in \mathcal{V}$ ) is also a bridge-addable class on the set of vertices  $V \setminus \cup_{i=1}^{k-1} V_i$ . It is worth to stress here that  $|V \setminus \cup_{i=1}^{k-1} V_i| \geq n - (k - 1)u$  is large enough (provided  $n$  is large enough), and thus, our previous results can be applied to these classes of graphs.

Consider the graphs in  $\mathcal{G}_n$  with  $k + 1$  components such that the  $k$  smallest ones are isomorphic to  $U_1, \dots, U_k$  and where one component isomorphic to  $U_k$  is marked. By counting these graphs in two ways, for  $n$  large enough, we have

$$m(U_1, \dots, U_k) \left| \mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}} \right| = \sum_{(V_1, \dots, V_{k-1}) \in \mathcal{V}} |\mathcal{H}^{2, U_k}(V_1, \dots, V_{k-1})| \quad (32)$$

Therefore,

$$\begin{aligned} & \frac{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n|} = \\ &= \frac{1}{m(U_1, \dots, U_k)} \sum_{(V_1, \dots, V_{k-1}) \in \mathcal{V}} \frac{|\mathcal{H}^{2, U_k}(V_1, \dots, V_{k-1})|}{|\mathcal{H}(V_1, \dots, V_{k-1})|} \cdot \frac{|\mathcal{H}(V_1, \dots, V_{k-1})|}{|\mathcal{H}^{(1)}(V_1, \dots, V_{k-1})|} \cdot \frac{|\mathcal{H}^{(1)}(V_1, \dots, V_{k-1})|}{|\mathcal{G}_n|}. \end{aligned} \quad (33)$$

Thus it suffices to estimate the three ratios in the sum above.

Let  $\theta_1 := \frac{\theta_k}{8}$  and  $\theta_{k-1} := \frac{\theta_k}{8}$ . Let  $\zeta_1$  be the constant obtained from Proposition 13 with  $\theta_1$  and  $U = U_k$ . Let  $\zeta_2$  be the constant obtained by induction with  $k-1$ ,  $\theta_{k-1}$  and  $U_1, \dots, U_{k-1}$ . We set  $\zeta_0 := \min \left\{ \zeta_1, \zeta_2, \frac{\theta_k}{20}, k^{-2} \right\}$ .

Let us first show that most of graphs in  $\mathcal{H}$  are in classes  $\mathcal{H}(V_1, \dots, V_{k-1})$  that are close to be tight. Let  $\mathcal{V}_0 \subset \mathcal{V}$  be the set of  $(k-1)$ -tuples such that  $\mathcal{H}(V_1, \dots, V_{k-1})$  satisfies

$$\Pr(H \in \mathcal{H}(V_1, \dots, V_{k-1}) \text{ connected}) \geq (1 + \zeta_0)e^{-1/2}, \quad (34)$$

and let  $\mathcal{H}_0 = \cup_{(V_1, \dots, V_{k-1}) \in \mathcal{V}_0} \mathcal{H}(V_1, \dots, V_{k-1})$ .

**Claim.** There exists a  $\zeta_3$  such that if  $\mathcal{G}$  is  $\zeta_3$ -tight and  $n$  is large enough, we have

$$|\mathcal{H}_0| \leq \zeta_0 |\mathcal{H}|.$$

*Proof of the Claim.* We first need to define the following subclasses that generalize  $\mathcal{H}(V_1, \dots, V_{k-1})$ . For any  $(k-1)$ -tuple of trees  $(W_1, W_2, \dots, W_{k-1})$ , we define

$$\mathcal{J}(W_1, \dots, W_{k-1}; V_1, \dots, V_{k-1}) = \{G[V \setminus \cup_{i=1}^{k-1} V_i] : G \in \mathcal{G}_n, G[V_1] \equiv W_1, \dots, G[V_{k-1}] \equiv W_{k-1}\}.$$

Similarly as in (31) to avoid problems of multiplicity, we define the following subsets that generalize  $\mathcal{V}$ ,

$$\mathcal{V}(W_1, \dots, W_{k-1}) = \{(V_1, \dots, V_{k-1}), V_i \subset [n] \text{ disjoint; if } W_i \equiv W_j \text{ then } V_i < V_j\}.$$

We stress here that, by definition, for any non-empty class  $\mathcal{J}(W_1, \dots, W_{k-1}; V_1, \dots, V_{k-1})$  such that  $\mathcal{H}(V_1, \dots, V_{k-1})$  is non-empty, we have  $|W_i| \leq u$ , for every  $1 \leq i \leq k-1$ . As before, we note that  $\mathcal{J}(W_1, \dots, W_{k-1}; V_1, \dots, V_{k-1})$  is bridge-addable. We will write  $\mathcal{J}(W_1, \dots, W_{k-1}) = \cup_{(V_1, \dots, V_{k-1}) \in \mathcal{V}(W_1, \dots, W_{k-1})} \mathcal{J}(W_1, \dots, W_{k-1}; V_1, \dots, V_{k-1})$  and we will write

$$\mathcal{J} = \bigcup_{\{W_1, \dots, W_{k-1}\}} \mathcal{J}(W_1, \dots, W_{k-1}),$$

where the union is taken over multisets of trees  $\{W_1, \dots, W_{k-1}\}$  and where for each multiset an arbitrary ordered tuple  $(W_1, \dots, W_{k-1})$  is chosen. Thus,  $\mathcal{J}$  can be understood as the set of graphs in  $\mathcal{G}_n$  with at least  $k$  components where exactly  $k-1$  of the non-largest ones are marked. In particular, we have:

$$|\mathcal{J}| = \sum_{j \geq 0} \binom{k+j-1}{k-1} |\mathcal{G}_n^{(k+j)}|. \quad (35)$$



Let  $\eta = \zeta_0^3$  and  $m$  such that  $\sum_{\ell \geq m-k} \frac{1}{\ell!} \leq \eta$  and  $m \geq k$ . By Lemma 5 there exists a  $\zeta_4$  such that if  $\mathcal{G}$  is  $\zeta_4$ -tight and  $n$  is large enough, we have for every  $i \leq m$

$$\frac{|\mathcal{G}_n^{(i)}|}{|\mathcal{G}_n|} = \frac{\left(\frac{1}{2} \pm \zeta_0^3\right)^{i-1}}{(i-1)!}.$$

Moreover, using the previous bound and (7), if  $i > m$ ,

$$\frac{|\mathcal{G}_n^{(i)}|}{|\mathcal{G}_n|} \leq \frac{\left(\frac{1}{2} + \zeta_0^3\right)^m}{(i-1)!}.$$

Therefore from (35) we obtain

$$\begin{aligned} |\mathcal{J}| &= \frac{\left(\frac{1}{2} \pm \zeta_0^3\right)^{k-1}}{(k-1)!} \left( \sum_{j=0}^{m-k} \frac{\left(\frac{1}{2} \pm \zeta_0^3\right)^j}{j!} \pm \left(\frac{1}{2} + \zeta_0^3\right)^{m-k+1} \sum_{j>m-k} \frac{1}{j!} \right) |\mathcal{G}_n| \\ &= \frac{\left(\frac{1}{2} \pm \zeta_0^3\right)^{k-1}}{(k-1)!} \left( e^{(1/2 \pm \zeta_0^3)} \pm 2\eta \right) |\mathcal{G}_n| \\ &= \left( 1 \pm \frac{\zeta_0^2}{10} \right) e^{1/2} |\mathcal{G}_n^{(k)}|, \end{aligned} \tag{36}$$

since  $\zeta_0 \leq k^{-2}$  and  $\zeta_0$  is a small constant.

Now we set  $\zeta_3 := \min \left\{ \frac{\zeta_0^2}{10(u^u)^k}, \zeta_4 \right\}$ . Fix  $W_1, \dots, W_{k-1}$ . Since  $\mathcal{J}(W_1, \dots, W_{k-1})$  is a disjoint union of bridge-addable classes  $(\mathcal{J}(W_1, \dots, W_{k-1}; V_1, \dots, V_{k-1}))$ , for each  $(V_1, \dots, V_{k-1})$  of graphs with  $n - \sum_{j=1}^{k-1} |W_j| \geq n - (k-1)u$  vertices, if  $n$  is large enough, by Theorem A applied to each class  $\mathcal{J}(W_1, \dots, W_{k-1}; V_1, \dots, V_{k-1})$ , we have

$$\begin{aligned} |\mathcal{J}(W_1, \dots, W_{k-1})| &\leq (1 + \zeta_3) e^{1/2} |\mathcal{G}_n^{k, \{W_1, \dots, W_{k-1}\}}| \\ &\leq \left( 1 + \frac{\zeta_0^2}{10(u^u)^k} \right) e^{1/2} |\mathcal{G}_n^{k, \{W_1, \dots, W_{k-1}\}}|. \end{aligned} \tag{37}$$

Since there are at most  $(u^u)^k$  multisets of unrooted trees  $\{W_1, \dots, W_k\}$  of order at most  $u$ , from (36) and (37), we have that for every  $W_1, \dots, W_{k-1}$ ,

$$|\mathcal{J}(W_1, \dots, W_{k-1})| \geq (1 - \zeta_0^2/5) e^{1/2} |\mathcal{G}_n^{k, \{W_1, \dots, W_{k-1}\}}|.$$

This holds in particular for  $\mathcal{H} = \mathcal{J}(U_1, \dots, U_{k-1})$ , implying

$$|\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}| \leq (1 + \zeta_0^2/4) e^{-1/2} |\mathcal{H}|, \tag{38}$$

since  $\zeta_0$  is a small constant.

For the sake of contradiction assume now that  $|\mathcal{H}_0| \geq \zeta_0 |\mathcal{H}|$ .

Since  $\mathcal{H} \setminus \mathcal{H}_0$  is a disjoint union of bridge-addable classes on  $n - \sum_{j=1}^{k-1} |U_j| \geq n - (k-1)u$  vertices, provided that  $n$  is large enough, Theorem A implies  $\mathbf{Pr}(H \in \mathcal{H} \setminus \mathcal{H}_0 \text{ connected}) \geq (1 - \zeta_3) e^{-1/2}$ .

Moreover, by definition of  $\mathcal{H}_0$ , we have  $\mathbf{Pr}(H \in \mathcal{H}_0 \text{ connected}) \geq (1 + \zeta_0)e^{-1/2}$ . We obtain

$$\begin{aligned}
|\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}| &= \mathbf{Pr}(H \in \mathcal{H} \text{ connected})|\mathcal{H}| \\
&= \mathbf{Pr}(H \in \mathcal{H} \setminus \mathcal{H}_0 \text{ connected})|\mathcal{H} \setminus \mathcal{H}_0| + \mathbf{Pr}(H \in \mathcal{H}_0 \text{ connected})|\mathcal{H}_0| \\
&\geq ((1 - \zeta_3)|\mathcal{H} \setminus \mathcal{H}_0| + (1 + \zeta_0)|\mathcal{H}_0|) e^{-1/2} \\
&\geq (1 + \zeta_0^2 - \zeta_3 + \zeta_0\zeta_3) e^{-1/2}|\mathcal{H}| \\
&\geq (1 + \zeta_0^2/2) e^{-1/2}|\mathcal{H}|,
\end{aligned}$$

which gives a contradiction with (38). This concludes the proof of the claim.  $\square$

We now set  $\zeta := \min\{\zeta_0, \zeta_3\}$ , where  $\zeta_3$  is the one given by the previous claim.

Let  $(V_1, \dots, V_{k-1}) \in \mathcal{V} \setminus \mathcal{V}_0$ ; that is, the class  $\mathcal{H}(V_1, \dots, V_{k-1})$  is  $\zeta_0$ -tight (and thus, also  $\zeta_1$ -tight). By Proposition 13 applied to the class  $\mathcal{H}(V_1, \dots, V_{k-1})$ , with the chosen  $\theta_1$  and  $U = U_k$ , and since the class is  $\zeta_1$ -tight and its elements have at least  $n - \sum_{j=1}^{k-1} |V_j| \geq n - (k-1)u$  vertices, we have

$$\frac{|\mathcal{H}^{2, U_k}(V_1, \dots, V_{k-1})|}{|\mathcal{H}(V_1, \dots, V_{k-1})|} = e^{-1/2} \frac{e^{-|U_k|}}{\text{Aut}_u(U_k)} \pm \frac{\theta_k}{8}. \quad (39)$$

Since  $\mathcal{H}(V_1, \dots, V_{k-1})$  is bridge-addable and since  $(V_1, \dots, V_{k-1}) \in \mathcal{V} \setminus \mathcal{V}_0$ , by Theorem A and by definition of  $\mathcal{V}_0$

$$\frac{|\mathcal{H}(V_1, \dots, V_{k-1})|}{|\mathcal{H}^{(1)}(V_1, \dots, V_{k-1})|} = e^{1/2}(1 \pm \zeta_0). \quad (40)$$

We proceed to bound the contribution of classes indexed by  $\mathcal{V}_0$ . Using again the previous claim,

$$\begin{aligned}
\sum_{(V_1, \dots, V_{k-1}) \in \mathcal{V}_0} |\mathcal{H}^{(1)}(V_1, \dots, V_{k-1})| &\leq |\mathcal{H}_0| \leq \zeta_0 |\mathcal{H}| \\
&\leq \zeta_0 (1 - \zeta_0)^{-1} |\mathcal{H} \setminus \mathcal{H}_0| \\
&\leq 2\zeta_0 \sum_{(V_1, \dots, V_{k-1}) \in \mathcal{V} \setminus \mathcal{V}_0} |\mathcal{H}^{(1)}(V_1, \dots, V_{k-1})|,
\end{aligned} \quad (41)$$

where the last inequality comes from (40) and the fact that  $\zeta_0$  is a small constant. Therefore,

$$\begin{aligned}
|\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}| &= \sum_{(V_1, \dots, V_{k-1}) \in \mathcal{V}} |\mathcal{H}^{(1)}(V_1, \dots, V_{k-1})| \\
&= (1 \pm 2\zeta_0) \sum_{(V_1, \dots, V_{k-1}) \in \mathcal{V} \setminus \mathcal{V}_0} |\mathcal{H}^{(1)}(V_1, \dots, V_{k-1})|.
\end{aligned}$$

Using the induction hypothesis for  $k-1$ , with the chosen  $\theta_{k-1}$  and  $U_1, \dots, U_{k-1}$ , and since  $\mathcal{G}$  is

$\zeta_2$ -tight and its elements have at least  $n - \sum_{j=1}^{k-1} |V_j| \geq n - (k-1)u$  vertices, it follows that

$$\begin{aligned}
\sum_{(V_1, \dots, V_{k-1}) \in \mathcal{V} \setminus \mathcal{V}_0} \frac{|\mathcal{H}^{(1)}(V_1, \dots, V_{k-1})|}{|\mathcal{G}_n|} &= (1 \pm 2\zeta_0)^{-1} \frac{|\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}|}{|\mathcal{G}_n|} \\
&= (1 \pm 2\zeta_0)^{-1} \left( e^{-1/2} \frac{e^{-\sum_{i=1}^{k-1} |U_i|}}{\text{Aut}_u(U_1, \dots, U_{k-1})} \pm \frac{\theta_k}{8} \right) \\
&= e^{-1/2} \frac{e^{-\sum_{i=1}^{k-1} |U_i|}}{\text{Aut}_u(U_1, \dots, U_{k-1})} \pm \frac{\theta_k}{4}. \tag{42}
\end{aligned}$$

We are now ready to estimate (33). We rewrite (33) as

$$\frac{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n|} = \frac{1}{m(U_1, \dots, U_k)} (\Sigma_{\mathcal{V}_0} + \Sigma_{\mathcal{V} \setminus \mathcal{V}_0}).$$

where  $\Sigma_{\mathcal{V}_0}$  and  $\Sigma_{\mathcal{V} \setminus \mathcal{V}_0}$  are the contribution to the sum of the elements indexed by  $\mathcal{V}_0$  and by  $\mathcal{V} \setminus \mathcal{V}_0$ , respectively.

In order to estimate  $\Sigma_{\mathcal{V}_0}$ , we note that  $|\mathcal{H}(V_1, \dots, V_{k-1})| \leq e|\mathcal{H}^{(1)}(V_1, \dots, V_{k-1})|$ , since the class  $\mathcal{H}(V_1, \dots, V_{k-1})$  is bridge-addable and using Theorem 2.5 in [MSW05]. Using (41), we obtain

$$\Sigma_{\mathcal{V}_0} \leq \frac{e\zeta_0|\mathcal{H}|}{|\mathcal{G}_n|} \leq 3\zeta_0 < \frac{\theta_k}{2}.$$

To estimate  $\Sigma_{\mathcal{V} \setminus \mathcal{V}_0}$ , we use (39), (40) and (42), to obtain that

$$\Sigma_{\mathcal{V} \setminus \mathcal{V}_0} = e^{-1/2} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_k) \text{Aut}_u(U_1, \dots, U_{k-1})} \pm \frac{\theta_k}{2}$$

Using the previous two estimates and (30), we get

$$\begin{aligned}
\frac{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n|} &= \frac{1}{m(U_1, \dots, U_k)} \cdot e^{-1/2} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_k) \text{Aut}_u(U_1, \dots, U_{k-1})} \pm \theta_k \\
&= e^{-1/2} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_1, \dots, U_k)} \pm \theta_k.
\end{aligned}$$

This concludes the proof of the first part.

**Proof of the second part.** We will prove the second part of the theorem, using the first part of it and by induction on  $k$ . For  $k = 1$ , the statement we want to prove is directly given by Proposition 13. Assume now that the statement is true for  $k - 1$ .

Set  $\hat{\theta}_k := e^{-1/2} \frac{e^{-ku}}{(ku)!} \theta_k$ . By the induction hypothesis, for  $\ell$ ,  $\theta_{k-1} := \frac{\hat{\theta}_k}{8}$ ,  $\delta_{k-1} := 2\delta$  and  $U_1, \dots, U_{k-1}$ , there exists a  $\zeta_{k-1}$  such that if  $n$  is large enough, we have

$$\frac{\sum_{\beta \notin \Xi(\delta_{k-1}, \ell)} |\mathcal{G}_{n, \beta}^{k, \{U_1, \dots, U_{k-1}\}}|}{|\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}|} < \frac{\hat{\theta}_k}{8}. \tag{43}$$

Since the first part of the theorem for  $k$  is already proved, we use it to estimate the ratio between  $\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}$  and  $\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}$ . For the first one we use the first part of the theorem for  $k$  with  $\theta_k := \frac{\hat{\theta}_k}{8}$  and  $U_1, \dots, U_k$  and the corresponding  $\zeta'_k$ . For the second one we use, as before, the first part of the theorem for  $k-1$  with  $\theta_{k-1}$  and  $U_1, \dots, U_{k-1}$  and the corresponding  $\zeta_{k-1}$ . Set  $\zeta := \min\{\zeta_{k-1}, \zeta'_k\}$  and let  $n$  be large enough.

Using (30), it follows that

$$\begin{aligned} \frac{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}|} &= \frac{e^{-1/2} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_1, \dots, U_k)} \pm \hat{\theta}_k/8}{e^{-1/2} \frac{e^{-\sum_{i=1}^{k-1} |U_i|}}{\text{Aut}_u(U_1, \dots, U_{k-1})} \pm \hat{\theta}_k/8} \\ &= \frac{e^{-|U_k|}}{m(U_1, \dots, U_{k-1}) \text{Aut}_u(U_k)} \left(1 \pm \frac{\theta_k}{3}\right). \end{aligned} \quad (44)$$

Let  $T_1, \dots, T_s$  be all the possible rooted versions of the unrooted tree  $U_k$ . Observe that  $|T_i| = |U_k|$  and that

$$\sum_{i=1}^s \frac{1}{\text{Aut}_r(T_i)} = \frac{|U_k|}{\text{Aut}_u(U_k)}. \quad (45)$$

Recall the definition of  $p(\mathcal{H}, T)$  given at the beginning of Section 3.4. We perform an exact double-counting argument between the graphs in  $\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}$  and in  $\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}$  using  $p(G, T_i)$  with  $G \in \mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}$ , similar to the one used in Section 2.1. In one direction, for any such graph  $G$ , we have exactly  $\sum_{i=1}^s \alpha^G(T_i) p(G, T_i)$  ways to construct a graph  $G' \in \mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}$  by removing an edge. In the other direction, there are exactly  $m(U_1, \dots, U_k) |U_k| (n - \sum_{i=1}^k |U_i|)$  ways to obtain a graph in  $\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}$  from one in  $\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}$  by adding an edge. Therefore, we have

$$\sum_{G \in \mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}} \sum_{i=1}^s \alpha^G(T_i) p(G, T_i) = m(U_1, \dots, U_k) |U_k| \left(n - \sum_{i=1}^k |U_i|\right) |\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}| \quad (46)$$

Using (44) and (45), it follows that

$$\begin{aligned} \frac{\sum_{G \in \mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}} \sum_{i=1}^s \alpha^G(T_i) p(G, T_i)}{n |\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}|} &= \frac{m(U_1, \dots, U_k) |U_k| \left(n - \sum_{i=1}^k |U_i|\right) |\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|}{n |\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}|} \\ &= \frac{n - \sum_{i=1}^k |U_i|}{n} \cdot \frac{|U_k| e^{-|U_k|}}{\text{Aut}_u(U_k)} \left(1 \pm \frac{\theta_k}{3}\right) \\ &= \sum_{i=1}^s \frac{e^{-|T_i|}}{\text{Aut}_r(T_i)} \left(1 \pm \frac{\theta_k}{2}\right), \end{aligned}$$

provided that  $n$  is large enough.

Since for every  $G \in \mathcal{G}_n$ ,  $\sum_{i=1}^s \alpha^G(T_i) p(G, T_i) \leq n$ , it follows that

$$\begin{aligned}
& \frac{\sum_{\beta \in \Xi(\delta_{k-1}, \ell)} \left( \sum_{i=1}^s \beta(T_i) p(\mathcal{G}_{n, \beta}^{k, \{U_1, \dots, U_{k-1}\}}, T_i) \right) \cdot |\mathcal{G}_{n, \beta}^{k, \{U_1, \dots, U_{k-1}\}}|}{n |\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}|} \\
&= \frac{\sum_{G \in \mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}} \sum_{i=1}^s \alpha^G(T_i) p(G, T_i)}{n |\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}|} \pm \frac{\sum_{\beta \notin \Xi(\delta_{k-1}, \ell)} |\mathcal{G}_{n, \beta}^{k, \{U_1, \dots, U_{k-1}\}}|}{|\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}|} \\
&= \sum_{i=1}^s \frac{e^{-|T_i|}}{\text{Aut}_r(T_i)} \left( 1 \pm \frac{5\theta_k}{8} \right). \tag{47}
\end{aligned}$$

If  $G'$  is obtained from  $G$  by removing an edge that creates a component isomorphic to  $U_k$ , then  $|\alpha^G(T) - \alpha^{G'}(T)| \leq |U_k| \leq u$  for every  $T \in \mathcal{T}$ . Therefore, if  $G \in \mathcal{G}_{n, \alpha}^{k, \{U_1, \dots, U_{k-1}\}}$  for some  $\beta \in \Xi(\delta_{k-1}, \ell)$ , then  $G' \in \mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}$  is such that  $\alpha^{G'} \in \Xi(\delta, \ell)$  (recall that  $\delta_{k-1} = \delta/2$ ), provided that  $n$  is large enough. We thus obtain a local version of (46)

$$\begin{aligned}
& \sum_{\beta \in \Xi(\delta_{k-1}, \ell)} \left( \sum_{i=1}^s \beta(T_i) p(\mathcal{G}_{n, \beta}^{k, \{U_1, \dots, U_{k-1}\}}, T_i) \right) \cdot |\mathcal{G}_{n, \alpha}^{k, \{U_1, \dots, U_{k-1}\}}| \\
& \leq m(U_1, \dots, U_k) |U_k| \left( n - \sum_{i=1}^k |U_j| \right) \sum_{\beta \in \Xi(\delta, \ell)} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}|
\end{aligned}$$

Using (47), the last inequality and (44), it follows that

$$\begin{aligned}
& \sum_{i=1}^s \frac{e^{-|T_i|}}{\text{Aut}_r(T_i)} \left( 1 - \frac{5\theta_k}{8} \right) \\
& \leq \frac{1}{n |\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}|} \sum_{\beta \in \Xi(\delta_{k-1}, \ell)} \left( \sum_{i=1}^s \beta(T_i) p(\mathcal{G}_{n, \beta}^{k, \{U_1, \dots, U_{k-1}\}}, T_i) \right) \cdot |\mathcal{G}_{n, \beta}^{k, \{U_1, \dots, U_{k-1}\}}| \\
& \leq \frac{e^{-|U_k|}}{\text{Aut}_u(U_k) |\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|} \sum_{\beta \in \Xi(\delta, \ell)} \frac{(n - \sum_{i=1}^k |U_j|) |U_k|}{n} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}| \left( 1 + \frac{\theta_k}{3} \right) \\
& \leq \frac{|U_k| e^{-|U_k|}}{\text{Aut}_u(U_k)} \cdot \frac{\sum_{\beta \in \Xi(\delta, \ell)} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|} \left( 1 + \frac{\theta_k}{3} \right) \\
& = \sum_{i=1}^s \frac{e^{-|T_i|}}{\text{Aut}_r(T_i)} \cdot \frac{\sum_{\beta \in \Xi(\delta, \ell)} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|} \left( 1 + \frac{\theta_k}{3} \right),
\end{aligned}$$

where we used (45) for the last equality. We conclude,

$$\frac{\sum_{\beta \in \Xi(\delta, \ell)} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|} \geq 1 - \theta_k,$$

which finishes the proof of the theorem.  $\square$

## 4 From classes of forests to classes of graphs

In this section we extend the results of the previous section (where we obtained Theorem 1 for classes of forests) to general bridge-addable classes, concluding the proof of Theorem 1. In 4.1 we prove that graphs in  $\zeta$ -tight bridge-addable classes tend to have many removable edges, and in 4.2 we use this property and the results of Section 3 to conclude the proof of Theorem 1. Finally in 4.3 we give the proof of Corollary 3.

### 4.1 Removable edges in tight bridge-addable classes of graphs

A *2-block* of a graph  $G$  is a maximal 2-edge-connected graph (we assume that the graph composed by a single vertex is also 2-edge-connected). Every graph admits a unique decomposition into 2-blocks, joined by edges in a tree-like fashion.

For a graph class  $\mathcal{G}_n$ , we can consider the coarsest partition

$$\mathcal{G}_n = \bigsqcup_i \mathcal{H}_n^{[i]} \quad (48)$$

into subclasses  $\mathcal{H}_n^{[1]}, \mathcal{H}_n^{[2]}, \dots$  such that every two graphs in the same subclass have the same 2-blocks. By construction, if  $\mathcal{G}_n$  is bridge-addable, then every subclass  $\mathcal{H}_n^{[i]}$  is also bridge-addable.

For each such subclass  $\mathcal{H}$ , we assume that we have chosen, arbitrarily and once and for all, a spanning tree for each 2-block of the graphs in  $\mathcal{H}$ . We denote by  $\mathcal{F}_{\mathcal{H}}$  the class of forests obtained by replacing each 2-block with the corresponding spanning tree in each graph in  $\mathcal{H}$ . This is well-defined, since, by construction, graphs in the same subclass have the same 2-blocks. Moreover, the class  $\mathcal{F}_{\mathcal{H}}$  is also bridge-addable and the component structure (number and size) of each graph  $H \in \mathcal{H}$  is preserved in the corresponding forest  $F_H \in \mathcal{F}_{\mathcal{H}}$ . This construction was introduced in [BBG08], to which we refer for more details.

We start with the following useful lemma that says, loosely speaking, that most graphs in a  $\zeta$ -tight belong to subclasses  $\mathcal{H}_n^{[i]}$  that are themselves close to be tight.

**Lemma 15.** *For every  $\zeta_0 > 0$  there exists  $\zeta > 0$  such that if  $n$  is large enough, for any bridge-addable class  $\mathcal{G}$  that is  $\zeta$ -tight, the following is true. Let  $\mathcal{H}_n^{[1]}, \mathcal{H}_n^{[2]}, \dots$  be the partition of  $\mathcal{G}_n$  in bridge-addable subclasses defined above and let  $S_n(\zeta_0)$  be the set of values  $i$  such that:*

$$\Pr(H_n \in \mathcal{H}_n^{[i]} \text{ connected}) \leq (1 + \zeta_0)e^{-1/2}, \quad (49)$$

where  $H_n \in \mathcal{H}_n^{[i]}$  denotes a uniform random graph in  $\mathcal{H}_n^{[i]}$ . Then we have

$$\left| \bigcup_{i \in S_n(\zeta_0)} \mathcal{H}_n^{[i]} \right| \geq (1 - \zeta_0)|\mathcal{G}_n|. \quad (50)$$

*Proof.* The proof is direct by an averaging argument in a similar way as in the claim inside the proof of Theorem 14.  $\square$

The next theorem states that  $\zeta$ -tight bridge-addable classes of graphs (not only forests) are, in fact, close to be bridge-alterable. In what follows, we say that a vertex  $v$  in  $G_n$  is *connected to the*

bulk of  $G_n$  through a cut-edge, if there is a cut-edge  $e$  incident to  $v$  such after removing  $e$ , the newly created component not containing  $v$  has size at least  $3n/4$ . Note that for each  $v$  there is at most one edge  $e$  with this property. The connected component containing  $v$  after removing  $e$  is called a *pendant graph*. The edge  $e$  can *a priori* be removable or not, and if it is we say that  $v$  is *connected to the bulk of  $G_n$  through a removable cut-edge*.

**Lemma 16.** *For every  $\theta$ , there exist a  $\zeta$  and an  $\ell$  such that provided that  $n$  is large enough, for every  $\zeta$ -tight bridge-addable class  $\mathcal{G}$ , we have that if  $G_n$  is a graph chosen uniformly at random in  $\mathcal{G}_n$ , and  $V_n$  is a vertex chosen uniformly at random in  $G_n$ , the following holds with probability at least  $1 - \theta$ :  $V_n$  is connected to the bulk of  $G_n$  through a removable cut-edge and the corresponding pendant graph has order at most  $\ell$ .*

*Proof.* We first prove the lemma for bridge-addable classes of forests and then we transfer it to general bridge-addable classes of graphs.

Assume that  $\mathcal{G}_n$  is composed by forests. We first show that there exists an  $\ell$  such that if  $G_n$  is a graph chosen uniformly at random from  $\mathcal{G}_n$ , then with probability at least  $(1 - \theta/4)$  we have that  $p(\mathcal{G}_n, T) \geq 1 - \theta/4$  for every  $T \in \mathcal{T}_{\leq \ell}$ . Then we will prove that with probability at least  $1 - \theta$ , most of the pendant trees in  $G_n$  have size at most  $\ell$ .

From Lemma 6, we can choose  $\ell$  large enough such that

$$\sum_{T \in \mathcal{T}_{\leq \ell}} \frac{e^{-|T|}}{\text{Aut}_r(T)} \geq 1 - \frac{\theta}{10},$$

$$e^{-1/2} \sum_{k=0}^{\ell} \sum_{\{U_1, \dots, U_k\} \in \mathcal{U}_{\leq \ell}} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_1, \dots, U_k)} \geq 1 - \frac{\theta}{10}.$$

Let  $T \in \mathcal{T}_{\leq \ell}$  be a given rooted tree, we will show that  $p(\mathcal{G}_n, T) \geq 1 - \theta/4$ . Let  $\lambda$  be the size of the equivalence class of the root of  $T$  (that is the number of vertices where  $T$  can be re-rooted giving rise to another copy of  $T$ ). For every  $k \leq \ell$  and every  $U_1, \dots, U_k$  of order at most  $\ell$  such that  $U_k$  is the unrooted version of  $T$ , we will write the ratio between  $|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|$  and  $|\mathcal{G}_n|$  in two ways. We select  $\zeta$  small enough and  $n$  large enough, such that we can apply Theorem 14 for every  $k \leq \ell$ , for  $\theta_k = \tilde{\theta}$  (to be fixed later) and for every  $U_1, \dots, U_k$  of size at most  $\ell$ . If  $\mathcal{G}$  is  $\zeta$ -tight, we obtain

$$\frac{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n|} = e^{-1/2} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_1, \dots, U_k)} \pm \tilde{\theta}.$$

As before, we perform an exact local double-counting argument with the difference that now we only count those graphs  $G' \in \mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}$  that can be obtained from  $G \in \mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}$  by removing an edge from where a copy of  $T$  is pending. This can only be done if  $U_k$  is the unrooted version of  $T$  and if the edge that connects  $T$  to the rest of  $G$  is removable. Moreover, if  $G'$  is obtained from  $G$  in such a way, for every  $T_0 \in \mathcal{T}_{\leq \ell}$  we have  $|\alpha^G(T_0) - \alpha^{G'}(T_0)| \leq |T| \leq \frac{\tilde{\theta}n}{2}$ . In one direction, given a graph  $G \in \mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}$  there are exactly  $p(G, T)\alpha^G(T)$  many such ways to obtain a graph in  $\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}$ , and in the other one, exactly  $\lambda m(U_1, \dots, U_k)(n - \sum_{j=1}^k |U_j|)$  many ones. Applying Theorem 14 twice with  $\theta_k = \tilde{\theta}$  and  $\delta = \tilde{\theta}/2$ , if  $\zeta$  is small enough and  $n$  is

large enough, then if  $\mathcal{G}$  is  $\zeta$ -tight, we obtain

$$\begin{aligned}
\frac{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n|} &\leq \frac{1}{|\mathcal{G}_n|} \sum_{\beta \in \Xi(\delta, \ell)} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}| (1 + \tilde{\theta}) \\
&\leq \frac{1}{|\mathcal{G}_n| m(U_1, \dots, U_k)} \sum_{\beta \in \Xi(\tilde{\theta}, \ell)} \sum_{G \in \mathcal{G}_{n, \beta}^{k, \{U_1, \dots, U_{k-1}\}}} \frac{p(G, T) \alpha^G(T)}{(n - \sum_{j=1}^{k-1} |U_j|) \lambda} (1 + \tilde{\theta}) \\
&= \frac{1}{|\mathcal{G}_n| m(U_1, \dots, U_k)} \sum_{\beta \in \Xi(\tilde{\theta}, \ell)} |\mathcal{G}_{n, \beta}^{k, \{U_1, \dots, U_{k-1}\}}| \frac{p(\mathcal{G}_{n, \beta}^{k, \{U_1, \dots, U_{k-1}\}}, T) \beta(T)}{(n - \sum_{j=1}^{k-1} |U_j|) \lambda} (1 + \tilde{\theta}) \\
&\leq \frac{1}{m(U_1, \dots, U_k)} \cdot \frac{|\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}|}{|\mathcal{G}_n|} \cdot \frac{p(\mathcal{G}'_n, T)}{\lambda} \left( \frac{e^{-|T|}}{\text{Aut}_r(T)} + \tilde{\theta} \right) (1 + \tilde{\theta}) \\
&\leq \frac{1}{m(U_1, \dots, U_k)} \left( e^{-1/2} \frac{e^{-\sum_{i=1}^{k-1} |U_i|}}{\text{Aut}_u(U_1, \dots, U_{k-1})} + \tilde{\theta} \right) p(\mathcal{G}'_n, T) \frac{e^{-|U_k|}}{\text{Aut}_u(U_k)} (1 + 3\tilde{\theta}) \\
&\leq e^{-1/2} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_1, \dots, U_k)} \cdot p(\mathcal{G}'_n, T) (1 + 5\tilde{\theta}) .
\end{aligned}$$

where  $\mathcal{G}'_n$  is the class formed by the union of  $\mathcal{G}_{n, \beta}^{k, \{U_1, \dots, U_{k-1}\}}$  for  $\beta \in \Xi(\delta, \ell)$ . In the previous inequalities we have used that  $\text{Aut}_u(U_k) = \lambda \text{Aut}_r(T)$  and (45).

Combining these two expressions and since  $|\mathcal{G}'_n| \geq (1 - \tilde{\theta}) |\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}|$  (by Theorem 14), we obtain that for every rooted tree  $T \in \mathcal{T}_{\leq \ell}$ ,

$$p(\mathcal{G}_n^{k, \{U_1, \dots, U_{k-1}\}}, T) \geq 1 - 8\tilde{\theta} . \quad (51)$$

Now we set  $\tilde{\theta} := \theta \ell^{-(\ell^2+1)}/10$ . Applying Theorem 14 for every  $k \leq \ell$ ,  $\theta_k = \tilde{\theta}$  and  $U_1, \dots, U_k$ , and using the definition of  $\ell$

$$\begin{aligned}
\sum_{k=0}^{\ell} \sum_{U_1, \dots, U_k \in \mathcal{U}_{\leq \ell}} \frac{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n|} &= e^{-1/2} \sum_{k=0}^{\ell} \sum_{U_1, \dots, U_k \in \mathcal{U}_{\leq \ell}} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_1, \dots, U_k)} - \tilde{\theta} \ell (\ell^\ell)^\ell \\
&\geq 1 - \frac{\theta}{5} , \quad (52)
\end{aligned}$$

By averaging (51) over all  $k$  and  $U_1, \dots, U_{k-1}$  and using the last equation, for every  $T \in \mathcal{T}_{\leq \ell}$ , we obtain

$$p(\mathcal{G}_n, T) \geq 1 - \theta/4 , \quad (53)$$

which proves the first part.

Let us now show that there are many removable edges that isolate a tree of size at most  $\ell$ . Choose  $G_n$  uniformly at random from  $\mathcal{G}_n$  and then choose  $V_n$  uniformly at random from  $V(G_n)$ . Let  $A_1$  be the event that  $V_n$  is connected to the bulk of  $G_n$  through a removable cut-edge and let  $A_2$  be the event that the pendant tree rooted at  $V_n$  has order at most  $\ell$ . We want to show that  $\Pr(A_1 \cap A_2) \geq 1 - \theta$ .



Again, by applying Theorem 14 for every  $k \leq \ell$ ,  $\theta_k = \tilde{\theta}$  and  $U_1, \dots, U_k$ , and using (52), we obtain

$$\begin{aligned} \sum_{\beta \in \Xi(\delta, \ell)} |\mathcal{G}_{n, \beta}| &\geq \sum_{k=0}^{\ell} \sum_{U_1, \dots, U_k \in \mathcal{U}_{\leq \ell}} \sum_{\beta \in \Xi(\delta, \ell)} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}| \\ &\geq \sum_{k=0}^{\ell} \sum_{U_1, \dots, U_k \in \mathcal{U}_{\leq \ell}} |\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}| - \tilde{\theta} \ell (\ell^\ell)^\ell \\ &\geq (1 - \theta/4) |\mathcal{G}_n|. \end{aligned}$$

Moreover, for every  $\beta \in \Xi(\delta, \ell)$  and by our choice of  $\ell$ , we have that  $\sum_{T \in \mathcal{T}_{\leq \ell}} \frac{\beta(T)}{n} \geq \sum_{T \in \mathcal{T}_{\leq \ell}} \frac{e^{|T|}}{\text{Aut}_r(T)} - \delta \ell^\ell \geq 1 - \theta/5$ . It follows that  $\mathbf{Pr}(A_2) \geq 1 - \theta/2$ .

Moreover, the probability of  $A_1$  given  $A_2$  can be written as a convex combination of  $p(\mathcal{G}_n, T)$  with  $T \in \mathcal{T}_{\leq \ell}$ . Therefore, by (53), we have that  $\mathbf{Pr}(A_1 | A_2) \geq 1 - \theta/4$ .

We conclude that

$$\mathbf{Pr}(A_1 \cap A_2) = 1 - \mathbf{Pr}(\overline{A_1} \cup \overline{A_2}) \geq 1 - (\mathbf{Pr}(\overline{A_2}) + \mathbf{Pr}(\overline{A_1} | A_2)) \geq 1 - 3\theta/4, \quad (54)$$

which concludes the proof of the theorem when all graphs in  $\mathcal{G}$  are forests.

In order to extend the result to general classes of graphs, we use the approach introduced in [BBG08]. For such a purpose, let  $\mathcal{G}_n$  be a general class of graphs.

Let  $\mathcal{H}_n^{[1]}, \mathcal{H}_n^{[2]}, \dots$  be the partition of  $\mathcal{G}_n$  into subclasses defined at the beginning of this section. Given  $\zeta_0$  (to be fixed later), we let  $S_n = S_n(\zeta_0)$  be the set of indices given by Lemma 15, and we fix an index  $i \in S_n$ . We let  $\mathcal{H} := \mathcal{H}_n^{[i]}$  be the corresponding subclass of  $\mathcal{G}_n$  and we let  $\mathcal{F}_{\mathcal{H}}$  be the corresponding class of forests. We observe that  $\mathcal{F}_{\mathcal{H}}$  is  $\zeta_0$ -tight and bridge-addable.

Since Lemma 16 holds for classes of forests, we can apply it to  $\mathcal{F}_{\mathcal{H}}$ . Note that if a cut-edge is removable for a forest  $F_H \in \mathcal{F}_{\mathcal{H}}$ , then the edge does not belong to any of the 2-blocks of the corresponding graph  $H \in \mathcal{H}$ . This implies that this cut-edge is also removable for  $H \in \mathcal{H}$ . Moreover, if its removal in  $F_H$  results in a tree of size at most  $\ell$ , then its removal in  $H$  results in a *graph* of size at most  $\ell$ . Therefore, the result obtained in (54) for  $\mathcal{F}_{\mathcal{H}}$  naturally transfers to the class  $\mathcal{H}$ , provided we change “trees” by “graphs” in what results after deleting a removable edge.

Moreover, if we choose  $\zeta_0$  small enough with respect to  $\theta$ , then there exists  $\zeta$  such that if  $\mathcal{G}$  is  $\zeta$ -tight and  $n$  is large enough, by (50), at least  $(1 - \theta/4)|\mathcal{G}_n|$  graphs in  $\mathcal{G}_n$  are in subclasses  $\mathcal{H}_n^{[i]}$  with  $i \in S_n$ , concluding that the lemma also holds for general classes of graphs  $\mathcal{G}_n$ .  $\square$

A direct consequence of the previous lemma is that, for most of the graphs in  $\mathcal{G}_n$ , most of their vertices are in 2-blocks of size 1. Indeed, if a graph with  $n$  vertices has  $(1 - \theta')n$  cut-edges for some  $\theta' > 0$ , it has at least  $(1 - 2\theta')n$  2-blocks of size one by an easy counting argument.

Our next goal is to use this observation to show that not only the pendant graphs obtained when deleting a removable edge have bounded size, as Lemma 16 ensures, but, in fact, they are pendant trees. For every class  $\mathcal{G}_n$  and every  $t \geq 1$ , if  $G_n$  is chosen uniformly at random from  $\mathcal{G}_n$  and  $V_n$  is chosen uniformly at random from  $V(G_n)$ , then let  $q(\mathcal{G}_n, t)$  be the probability that  $V_n$  is connected to the bulk of  $G_n$  through a removable cut-edge and the corresponding pendant graph is a tree of order at most  $t$ . Observe that if  $\mathcal{G}$  is a subclass of forests, Lemma 16 implies that for every  $\theta$ , and under some conditions, there exists an  $\ell$  such that  $q(\mathcal{G}_n, \ell) \geq 1 - \theta$ . Next lemma shows that the same holds for general classes of graphs.

**Lemma 17.** *For every  $\vartheta$ , there exist a  $\zeta$  and a  $t$ , such that if  $\mathcal{G}$  is a  $\zeta$ -tight bridge-addable class and  $n$  is large enough, then*

$$q(\mathcal{G}_n, t) \geq 1 - \vartheta.$$

*Proof.* Given  $G \in \mathcal{G}_n$  and a vertex  $v \in V(G)$  that is connected to the bulk of  $G$  through a cut-edge  $e$ , we denote by  $X_G(v)$  the pendant graph (containing  $v$ ) obtained when deleting  $e$  from  $G$ . Given  $G_n$  chosen uniformly at random from  $\mathcal{G}_n$  and  $V_n$  chosen uniformly at random from  $V(G_n)$ , as before, we define  $A_1$  as the event that  $V_n$  is connected to the bulk of  $G_n$  through a removable cut-edge and  $A_2$  as the event that  $X_{G_n}(V_n)$  has order at most  $t$ . Also, let  $A_3$  be the event that  $X_{G_n}(V_n)$  is a tree. It is implicit in the definition of  $A_2$  and  $A_3$  that  $V_n$  should be connected to the bulk of  $G_n$  through a cut-edge, so in particular  $X_{G_n}(V_n)$  is well-defined. Note that:

$$\begin{aligned} q(\mathcal{G}_n, t) &= \Pr(A_1 \cap A_2 \cap A_3) = \Pr(A_1 \cap A_2) - \Pr(A_1 \cap A_2 \cap \overline{A_3}) \\ &\geq \Pr(A_1 \cap A_2) - \Pr(A_2 \cap \overline{A_3}), \end{aligned} \quad (55)$$

so we will proceed by bounding the last two probabilities.

Here we consider again the partition of  $\mathcal{G}_n$  into subclasses  $\mathcal{H}_n^{[1]}, \mathcal{H}_n^{[2]}, \dots$  defined above. Given  $\zeta_0$  (to be fixed later), there exists a  $\zeta$  such that for every  $\zeta$ -tight class  $\mathcal{G}$ , if  $n$  is large enough, we can consider  $S_n = S_n(\zeta_0)$  to be the set of indices given by Lemma 15. We let  $\mathcal{H} := \mathcal{H}_n^{[i]}$  be the corresponding subclass of  $\mathcal{G}_n$ , for some  $i \in S_n$ , and  $\mathcal{F}_{\mathcal{H}}$  be corresponding class of forests.

By Lemma 16 with  $\theta = \vartheta/3$ , if  $\zeta_0$  is small enough and,  $n$  and  $t$  are large enough, since  $\mathcal{H}$  is a  $\zeta_0$ -tight bridge-addable class of graphs with  $n$  vertices, then the probability that a uniformly chosen vertex  $W_n$  from a uniformly chosen forest  $F_n$  in  $\mathcal{F}_{\mathcal{H}}$  connects to the bulk of  $F_n$  through a removable cut-edge and that  $X_{F_n}(W_n)$  is a tree of order at most  $t$ , is at least  $1 - \vartheta/3$ . If this is the case, as we argued before, this edge is also a removable cut-edge in the graph in  $\mathcal{H}$  that corresponds to  $F_n$ . Thus, using (50) and provided that  $\zeta_0$  is small enough with respect to  $\vartheta$ , we can lower bound the first probability in (55) as follows

$$\Pr(A_1 \cap A_2) \geq 1 - \frac{\vartheta}{3} - \zeta_0 \geq 1 - \frac{\vartheta}{2}. \quad (56)$$

It remains to obtain an upper bound on  $\Pr(A_2 \cap \overline{A_3})$ . For this we first observe that, by using Lemma 16 again with  $\theta_2 = \frac{\vartheta}{7t}$ , and if  $\zeta_0$  is small enough, and,  $n$  and  $\ell$  are large enough, since  $\mathcal{H}$  is  $\zeta_0$ -tight, then the probability that a uniformly chosen vertex  $W_n$  from a uniformly chosen forest  $F_n$  in  $\mathcal{F}_{\mathcal{H}}$  is connected to the bulk of  $F_n$  through a removable cut-edge, is at least  $1 - \frac{\vartheta}{7t}$ . Using (50) again and provided that  $\zeta_0$  is small enough with respect to  $\vartheta$  and  $t$ , we obtain

$$\Pr(\overline{A_1}) \leq \frac{\vartheta}{7t} + \zeta_0 \leq \frac{\vartheta}{6t}. \quad (57)$$

We claim that

$$\Pr(A_2 \cap \overline{A_3}) \leq t \Pr(\overline{A_1}). \quad (58)$$

Assuming that (58) holds, together with (56) and with (57), we can now control each term in (55) to obtain

$$q(\mathcal{G}_n, t) \geq 1 - \frac{\vartheta}{2} - \frac{\vartheta}{6} \geq 1 - \vartheta.$$

Thus, it only remains to prove (58). For this we observe that if  $A_2 \cap \overline{A_3}$  holds, then  $X_{G_n}(V_n)$  contains at least one vertex  $V'_n$  which is not connected to the bulk of  $G_n$  through a cut-edge (since  $X_{G_n}(V_n)$  is a well-defined pendant graph, but it is not a tree). Moreover since  $A_2$  holds, the graph distance between  $V_n$  and  $V'_n$  is less than  $t$ . Conversely, it is easy to see that given any vertex  $v'$ , there are at most  $t$  vertices  $v$  at distance at less than  $t$  from  $v'$  that are connected to the bulk of  $G_n$  through a cut-edge and such that  $X_{G_n}(v)$  contains  $v'$ . The inequality (58) thus follows by double-counting such pairs of vertices.  $\square$

## 4.2 Proof of Theorem 1

We finally present the proof of our main theorem.

*Proof of Theorem 1.* Let us first prove *i*). We will first prove that for every  $k$ , every  $\theta$  and every  $U_1, \dots, U_k$ , and if  $\zeta$  is small enough and  $n$  large enough, then for every  $\zeta$ -tight bridge-addable class  $\mathcal{G}$ , we have

$$\left| \frac{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n|} - e^{-1/2} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_1, \dots, U_k)} \right| < \theta. \quad (59)$$

As before we consider the partition of  $\mathcal{G}_n$  into subclasses  $\mathcal{H}_n^{[1]}, \mathcal{H}_n^{[2]}, \dots$ . Given  $\zeta_0$  (to be fixed later), there exists a  $\zeta$  such that for every  $\zeta$ -tight class  $\mathcal{G}$ , if  $n$  is large enough, we can consider the set  $S_n = S_n(\zeta_0)$  given by Lemma 15.

Let  $\mathcal{H} := \mathcal{H}_n^{[i]}$  for  $i \in S_n$  and let  $\mathcal{F}_{\mathcal{H}}$  be the corresponding  $\zeta_0$ -tight class of forests. We can apply Theorem 14 for the given  $k$ ,  $\theta_k = \frac{\theta}{4}$ , and the given  $U_1, \dots, U_k$ . If  $\zeta_0$  is small enough and  $n$  is large enough, and since  $\mathcal{F}_{\mathcal{H}}$  is  $\zeta_0$ -tight, (59) holds for  $\mathcal{F}_{\mathcal{H}}$ .

It follows that

$$\begin{aligned} \left| \mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}} \right| &= \sum_{j \in S_n} \left| (\mathcal{H}_n^{[j]})^{k+1, \{U_1, \dots, U_k\}} \right| \pm \zeta_0 |\mathcal{G}_n| \\ &= \sum_{j \in S_n} \left| \mathcal{F}_{\mathcal{H}_n^{[j]}}^{k+1, \{U_1, \dots, U_k\}} \right| \pm \zeta_0 |\mathcal{G}_n| \\ &= \left( e^{-1/2} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_1, \dots, U_k)} \pm \theta_k \right) \sum_{j \in S_n} |\mathcal{F}_{\mathcal{H}_n^{[j]}}| \pm \zeta_0 |\mathcal{G}_n| \\ &= \left( e^{-1/2} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_1, \dots, U_k)} \pm \theta_k \right) (1 \pm \zeta_0) |\mathcal{G}_n| \pm \zeta_0 |\mathcal{G}_n| \\ &= \left( e^{-1/2} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_1, \dots, U_k)} \pm \theta \right) |\mathcal{G}_n|, \end{aligned}$$

provided that  $\zeta_0$  is small enough with respect to  $\theta$ . This proves (59).

To prove the first part of the theorem, let  $k_*$  be large enough such that

$$e^{-1/2} \sum_{k=0}^{k_*} \sum_{\{U_1, \dots, U_k\} \in \mathcal{U}_{\leq k_*}} \frac{e^{-\sum_{i=1}^k |U_i|}}{\text{Aut}_u(U_1, \dots, U_k)} \geq 1 - \frac{\epsilon}{4}. \quad (60)$$

The existence of such a  $k_*$  is, again, guaranteed by Lemma 6.

If  $\mathbf{f}$  is an unrooted unlabeled forest composed by trees  $U_1, \dots, U_k$ , then

$$\Pr(\text{Small}(G_n) \equiv \mathbf{f}) = \frac{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n|}.$$

We choose  $\theta := \epsilon k_*^{-k_*^2}/2$ .

Let  $\mathbf{f}_1$  be a forest composed by at most  $k_*$  trees of size at most  $k_*$ , then (59) gives that  $|\Pr(\text{Small}(G_n) \equiv \mathbf{f}_1) - p_\infty(\mathbf{f}_1)| < \epsilon$ .

Let  $\mathbf{f}_2$  be a forest with either more than  $k_*$  trees or where at least one of the trees has size larger than  $k_*$ . Since  $p_\infty$  is a probability distribution, by (60) we have  $p_\infty(\mathbf{f}_2) \leq \epsilon/4$ . Since  $\sum_{\mathbf{f}} \Pr(\text{Small}(G_n) \equiv \mathbf{f}) = 1$ , using again (59) and (60), we have

$$\begin{aligned} |\Pr(\text{Small}(G_n) \equiv \mathbf{f}_2) - p_\infty(\mathbf{f}_2)| &\leq \Pr(\text{Small}(G_n) \equiv \mathbf{f}_2) + p_\infty(\mathbf{f}_2) \\ &\leq 1 - \sum_{k=0}^{k_*} \sum_{\{U_1, \dots, U_k\} \in \mathcal{U}_{\leq k_*}} \frac{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n|} + p_\infty(\mathbf{f}_2) \\ &\leq \epsilon/4 + \theta k_*^{k_*^2} + \epsilon/4 = \epsilon \end{aligned}$$

This concludes the proof of *i*).

Let us now prove the following property from which *ii*) follows directly.

*iii) for every  $\epsilon, \eta$ , there exists  $\zeta$  such that for every  $\zeta$ -tight bridge-addable class  $\mathcal{G}$  and every  $n$  large enough, if  $\mathbf{f}$  is a fixed unrooted unlabeled forest, we have:*

$$\left| \Pr \left( \text{Small}(G_n) \equiv \mathbf{f}; \forall T \in \mathcal{T} : \left| \frac{\alpha_{G_n}(T)}{n} - a_\infty(T) \right| < \eta \right) - p_\infty(\mathbf{f}) \right| < \epsilon.$$

Similarly as before, we first prove that for every  $\theta, \eta, k, \ell$  and  $U_1, \dots, U_k$ , and provided that  $\zeta$  is small enough and  $n$  large enough, we have

$$\frac{\sum_{\beta \in \Xi(\eta, \ell)} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|} \geq 1 - \theta. \quad (61)$$

Recall the partition of  $\mathcal{G}_n$  into subclasses  $\mathcal{H}_n^{[1]}, \mathcal{H}_n^{[2]}, \dots$ . As before, let  $\mathcal{H} := \mathcal{H}_n^{[i]}$  for  $i \in S_n$  and let  $\mathcal{F}_\mathcal{H}$  be the corresponding class of forests. If we apply the second part of Theorem 14 with  $\theta_k = \theta/4$  and  $\delta = \eta/2$  to the class  $\mathcal{F}_\mathcal{H}$ , if  $\zeta_0$  is small enough and  $n$  large enough, then at least  $(1 - \theta/2)|\mathcal{H}^{k+1, \{U_1, \dots, U_k\}}|$  graphs  $G \in \mathcal{H}^{k+1, \{U_1, \dots, U_k\}}$  satisfy  $\alpha^{F_G} \in \Xi(\delta, \ell)$ .

Theorem 14 also shows that there exists a constant  $c_1 > 0$  such that  $|\mathcal{H}^{k+1, \{U_1, \dots, U_k\}}| \geq c_1 |\mathcal{H}|$ . By Lemma 17 with  $\vartheta := c_1 \min\{\theta/4, \delta\}$ , if  $\zeta_0$  is small enough and  $n$  large enough there exists a  $t$  such that with probability at least  $1 - \vartheta$ , a random vertex in a random graph of  $\mathcal{H}$  is connected via a removable cut-edge and the corresponding pending graph is a tree of order at most  $t$ . We can choose  $t \geq \ell$ . (Note that by doing so, we only increase the former probability.)

Therefore, if  $H_n$  is a random graph in  $\mathcal{H}^{k+1, \{U_1, \dots, U_k\}}$ , with probability at least  $1 - \theta/2 - \vartheta/c_1 > 1 - 3\theta/4$ , for every  $T \in \mathcal{T}_{\leq \ell}$ ,

$$\frac{\alpha^{H_n}(T)}{n} = \frac{e^{-|T|}}{\text{Aut}_r(T)} \pm \delta \pm \vartheta = \frac{e^{-|T|}}{\text{Aut}_r(T)} \pm 2\delta.$$

In other words, with probability at least  $1 - 3\theta/4$ , we have  $\alpha^{H_n} \in \Xi(2\delta, \ell) = \Xi(\eta, \ell)$ .

By *i*), we have that  $|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}| \geq c_2 |\mathcal{G}_n|$ , for some constant  $c_2 > 0$ . Therefore, there are at most  $\frac{\zeta_0}{c_2} |\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|$  graphs in classes  $\mathcal{H}_n^{[i]}$  that are not  $\zeta_0$ -tight. We conclude that, provided  $\zeta_0$  is small enough, the probability that a graph  $G'_n$  chosen at random from  $\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}$  satisfies  $\alpha^{G'_n} \in \Xi(\eta, \ell)$ , is at least  $1 - 3\theta/4 - \zeta_0/c_2 > 1 - \theta$ . This proves (61).

Let  $A(k, \nu)$  the event that for every  $T \in \mathcal{T}_{\leq k}$  we have  $\left| \frac{\alpha^{G_n(T)}}{n} - a_\infty(T) \right| < \nu$  (we might write  $k = \infty$  where  $\mathcal{T}_{\leq \infty} = \mathcal{T}$ ).

Since we have already proved *i*), we have that for every unrooted unlabeled forest  $\mathbf{f}$  with small components  $U_1, \dots, U_k$ , then

$$\mathbf{Pr}(A(\infty, \eta), \text{Small}(G_n) \equiv \mathbf{f}) \leq \mathbf{Pr}(\text{Small}(G_n) \equiv \mathbf{f}) \leq p_\infty(\mathbf{f}) + \epsilon. \quad (62)$$

By Lemma 6, if  $k_*$  is large enough, then

$$\sum_{T \in \mathcal{T}_{\leq k_*}} \frac{e^{-|T|}}{\text{Aut}_r(T)} > 1 - \frac{\eta}{4}.$$

Let  $T' \notin \mathcal{T}_{\leq k_*}$  and choose  $\rho = \eta k_*^{-k_*}/4$ . As before, by the properties of  $k_*$  we have that  $a_\infty(T') \leq \eta/4$  and, conditional on  $A(k_*, \rho)$ ,  $\frac{\alpha^{G_n(T')}}{n} \leq \eta/4 + \rho k_*^{k_*} = \eta/2$ . This implies that, conditional on  $A(k_*, \rho)$ , then  $A(k_*, \eta)$  implies  $A(\infty, \eta)$ .

If  $\mathbf{f}$  is an unrooted unlabeled forest with small components  $U_1, \dots, U_k$ , using (61), we have that for every  $\theta$ ,

$$\begin{aligned} & P(A(\infty, \eta) \mid \text{Small}(G_n) \equiv \mathbf{f}) \\ & \geq P(A(\infty, \eta) \mid \text{Small}(G_n) \equiv \mathbf{f}, A(k_*, \rho)) \cdot P(A(k_*, \rho) \mid \text{Small}(G_n) \equiv \mathbf{f}) \\ & \geq P(A(k_*, \eta) \mid \text{Small}(G_n) \equiv \mathbf{f}, A(k_*, \rho)) \cdot P(A(k_*, \rho) \mid \text{Small}(G_n) \equiv \mathbf{f}) \\ & = \frac{\sum_{\beta \in \Xi(\eta, k_*)} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}|}{\sum_{\beta \in \Xi(\rho, k_*)} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}|} \cdot \frac{\sum_{\beta \in \Xi(\rho, k_*)} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|} \\ & = \frac{\sum_{\beta \in \Xi(\eta, k_*)} |\mathcal{G}_{n, \beta}^{k+1, \{U_1, \dots, U_k\}}|}{|\mathcal{G}_n^{k+1, \{U_1, \dots, U_k\}}|} \\ & \geq 1 - \theta \end{aligned}$$

By *i*), we may assume that  $\mathbf{Pr}(\text{Small}(G_n) \equiv \mathbf{f}) \geq p_\infty(\mathbf{f}) - \epsilon/2$ . Choosing  $\theta := \epsilon/2$ , we conclude

$$\begin{aligned} P(A(\infty, \eta), \text{Small}(G_n) \equiv \mathbf{f}) &= P(A(\infty, \eta) \mid \text{Small}(G_n) \equiv \mathbf{f}) P(\text{Small}(G_n) \equiv \mathbf{f}) \\ &\geq (1 - \theta)(p_\infty(\mathbf{f}) - \epsilon/2) \geq p_\infty(\mathbf{f}) - \epsilon. \end{aligned}$$

Together with (62), this proves *iii*), which directly proves *ii*).  $\square$

### 4.3 Proof of Corollary 3

Corollary 3 is a simple consequence of our main result. Here we present a detailed proof.

*Proof of Corollary 3.* Let  $G$  be a graph with  $n$  vertices, let  $v \in V(G)$  and  $r \geq 1$ . The *ball* of radius  $r$  centered at  $v$ ,  $B_{G,r}(v)$ , is the graph induced in  $G$  by all vertices at distance at most  $r$  from  $v$ . The *hull* of radius  $r$ ,  $H_{G,r}(v)$  is the union of  $B_{G,r}(v)$  with all the connected components of  $G \setminus B_{G,r}(v)$  that are of size smaller than  $\frac{n}{3}$ , but are not components of  $G$ . We view the hull  $H_{G,r}(v)$  as a graph with a root (the vertex  $v$ ) and a set, possibly empty, of *exit vertices* (the vertices to which component(s) of size larger than  $\frac{n}{3}$  are attached). Note that the exit vertices are necessarily at distance  $r$  from the root. We extend the definition of hulls to infinite graphs, by replacing the condition “size smaller than  $\frac{n}{3}$ ” by the condition “finite size”.

For  $k \geq 0$ , let  $\mathcal{T}_{r,k}$  be the set of (unlabeled) trees with a marked root, and  $k$  marked distinct vertices at distance  $r$  from the root (exit vertices). For  $T \in \mathcal{T}_{r,k}$  and a rooted graph  $(G, v)$ , we write  $H_{G,r}(v) \equiv T$  if the hull  $H_{G,r}(v)$  is isomorphic to  $T$  as an unlabeled graph, where the isomorphism preserves the root and the exit vertices (in particular this implies that  $H_{G,r}(v)$  has  $k$  exit vertices). Then it is easy to see from the definition of  $(F_\infty, V_\infty)$  that we have, for any  $r \geq 1, k \geq 0$  and  $T \in \mathcal{T}_{r,k}$ :

$$\mathbf{Pr}(H_{F_\infty,r}(V_\infty) \equiv T) = q_\infty(T),$$

where

$$q_\infty(T) := \begin{cases} \frac{1}{\text{Aut}_{\text{path}}(T)} e^{-|T|} & \text{if } k = 1 \\ 0 & \text{if } k \neq 1, \end{cases}$$

where  $\text{Aut}_{\text{path}}(T)$  is the number of automorphisms of  $T$  preserving the path from the root to the exit vertex. Moreover, for any  $r \geq 1$  we have

$$\sum_{k \geq 0} \sum_{T \in \mathcal{T}_{r,k}} q_\infty(T) = 1. \quad (63)$$

Let  $r \geq 1$  and fix  $T \in \mathcal{T}_{r,1}$ , with root  $u$  and exit vertex  $w$ . Let  $T'$  be the element of  $\mathcal{T}$  obtained by re-rooting the tree  $T$  at  $w$ , and let  $m$  be the number of copies of the vertex  $u$  in  $T'$ . Then, clearly, there are *at least*  $m\alpha^G(T')$  vertices  $v$  in  $V(G)$  such that  $H_{G,r}(v) \equiv T$ . We thus have,

$$\mathbf{Pr}(H_{G,r}(V) \equiv T) \geq \frac{m\alpha^G(T')}{n},$$

where  $V$  is a uniform random vertex in  $G$ . Now let  $\mathcal{G}$  be a tight bridge-addable graph class, and, for every  $n \geq 1$ , let  $G_n$  be a uniform graph in  $\mathcal{G}_n$  and let  $V_n$  be a uniform vertex in  $G_n$ . By averaging over graphs in  $\mathcal{G}_n$  and using the second part of Theorem 2 we obtain:

$$\liminf_n \mathbf{Pr}(H_{G_n,r}(V_n) \equiv T) \geq \liminf_n \mathbf{E} \left( \frac{m\alpha^{G_n}(T')}{n} \right) \geq ma_\infty(T') = q_\infty(T) \quad (64)$$

where for the last equality we used  $m\text{Aut}_{\text{path}}(T) = \text{Aut}_r(T')$ . Now since the events  $H_{G_n,r}(V_n) \equiv T$  for  $T \in \cup_{k \geq 0} \mathcal{T}_{r,k}$  are disjoint, we have:

$$\sum_{k \geq 0} \sum_{T \in \mathcal{T}_{r,k}} \mathbf{Pr}(H_{G_n,r}(V_n) \equiv T) \leq 1.$$

From (63) and (64) we thus get that, for any  $r, k$  and  $T \in \mathcal{T}_{r,k}$  we have:

$$\lim_n \mathbf{Pr}(H_{G_n,r}(V_n) \equiv T) = q_\infty(T). \quad (65)$$

The last equation implies that, for any rooted graph  $B_0$  of radius  $r$  (where the radius is the greatest distance from a vertex to the root), we have:

$$\lim_n \mathbf{Pr}(B_{G_n,r}(V_n) \equiv B_0) = \mathbf{Pr}(B_{F_\infty,r}(V_\infty) \equiv B_0). \quad (66)$$

To see this, note that for every rooted graph  $B$  we have:

$$\mathbf{Pr}(B_{F_\infty,r}(V_\infty) \equiv B) = \sum_{k \geq 0} \sum_{\substack{T \in \mathcal{T}_{r,k} \\ T \triangleright B}} \mathbf{Pr}(H_{F_\infty,r}(V_\infty) \equiv T),$$

where  $T \triangleright B$  means that  $B_{T,r}(v) \equiv B$ , where  $v$  is the root of  $T$ .

It follows from this equality that for any  $B$ , any  $r \geq 1$  and any  $\epsilon$ , we can choose a finite subset  $\mathcal{T}' \subset \cup_{k \geq 0} \mathcal{T}_{r,k}$  such that  $\sum_{T \in \mathcal{T}', T \triangleright B} \mathbf{Pr}(H_{F_\infty,r}(V_\infty) \equiv T) \geq \mathbf{Pr}(B_{F_\infty,r}(V_\infty) \equiv B) - \epsilon$ . Using (65), it follows that

$$\begin{aligned} \liminf_n \mathbf{Pr}(B_{G_n,r}(V_n) \equiv B) &\geq \liminf_n \sum_{T \in \mathcal{T}', T \triangleright B} \mathbf{Pr}(H_{G_n,r}(V_n) \equiv T) \\ &\geq \sum_{T \in \mathcal{T}', T \triangleright B} \mathbf{Pr}(H_{F_\infty,r}(V_\infty) \equiv T) \\ &\geq \mathbf{Pr}(B_{F_\infty,r}(V_\infty) \equiv B) - \epsilon. \end{aligned}$$

Since this is true for any  $\epsilon > 0$  we thus have proved

$$\liminf_n \mathbf{Pr}(B_{G_n,r}(V_n) \equiv B) \geq \mathbf{Pr}(B_{F_\infty,r}(V_\infty) \equiv B). \quad (67)$$

It follows that

$$1 \geq \liminf_n \sum_B \mathbf{Pr}(B_{G_n,r}(V_n) \equiv B) \geq \sum_B \mathbf{Pr}(B_{F_\infty,r}(V_\infty) \equiv B) = 1,$$

where the sums are taken over all rooted graphs  $B$  of radius  $r$ , and using (67), Equation (66) holds for every  $B_0$ . This concludes the proof.  $\square$

## References

- [ABMR12] Louigi Addario-Berry, Colin McDiarmid, and Bruce Reed. Connectivity for bridge-addable monotone graph classes. *Combin. Probab. Comput.*, 21(6):803–815, 2012.
- [Ald93] David Aldous. The continuum random tree iii. *The Annals of Probability*, pages 248–289, 1993.
- [Ald98] David Aldous. Tree-valued Markov chains and Poisson-Galton-Watson distributions. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, volume 41 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 1–20. Amer. Math. Soc., Providence, RI, 1998.
- [BBG08] Paul Balister, Béla Bollobás, and Stefanie Gerke. Connectivity of addable graph classes. *J. Combin. Theory Ser. B*, 98(3):577–584, 2008.

- [BLL98] François Bergeron, Gilbert Labelle, and Pierre Leroux. *Combinatorial species and tree-like structures*, volume 67 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1998. Translated from the 1994 French original by Margaret Readdy, With a foreword by Gian-Carlo Rota.
- [BS01] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13 pp. (electronic), 2001.
- [CP15] Guillaume Chapuy and Guillem Perarnau. Connectivity in bridge-addable graph classes: the McDiarmid-Steger-Welsh conjecture. *Extended abstract in the proceedings of SODA 2016. Long version submitted for publication, see arXiv:1504.06344.*, 2015.
- [Erd66] Paul Erdős. On some new inequalities concerning extremal properties of graphs. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 77–81, 1966.
- [Erd67] Paul Erdős. Some recent results on extremal problems in graph theory. *Results, Theory of Graphs (Internat. Sympos., Rome, 1966)*, Gordon and Breach, New York, pages 117–123, 1967.
- [ES66] Paul Erdős and M Simonovits. A limit theorem in graph theory. In *Studia Sci. Math. Hung.* Citeseer, 1966.
- [FS09] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.
- [KP13] Mihyun Kang and Konstantinos Panagiotou. On the connectivity of random graphs from addable classes. *J. Combin. Theory Ser. B*, 103(2):306–312, 2013.
- [Lov12] László Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.
- [MSW05] Colin McDiarmid, Angelika Steger, and Dominic J. A. Welsh. Random planar graphs. *J. Combin. Theory Ser. B*, 93(2):187–205, 2005.
- [MSW06] Colin McDiarmid, Angelika Steger, and Dominic J. A. Welsh. Random graphs from planar and other addable classes. In *Topics in discrete mathematics*, volume 26 of *Algorithms Combin.*, pages 231–246. Springer, Berlin, 2006.
- [Rén59] Alfréd Rényi. Some remarks on the theory of trees. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 4:73–85, 1959.
- [Sim68] Miklós Simonovits. A method for solving extremal problems in graph theory, stability problems. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 279–319, 1968.



## A Generating functions of trees and forests

The purpose of this appendix is to recall some basic properties of uniform random forests and their generating functions, in particular proving Theorem B of the introduction and Lemma 6 of Section 2.4. The material of Section A.1 is a standard exercise in analytic combinatorics (for background, see the book [FS09]) but we are not aware of a reference containing precisely these results, so we prefer to recall them here. Section A.2 gives details about the example from Remark 3 in the introduction.

### A.1 Small components and pendant trees in uniform forests

For  $n \geq 1$  we let  $u_n, f_n$  be respectively the number of labeled trees and labeled forests on  $n$  vertices. Thus  $t_n := nu_n$  is the number of labeled rooted trees on  $n$  vertices.

We let  $T(z) = \sum_{n \geq 1} \frac{t_n}{n!} z^n$ ,  $U(z) = \sum_{n \geq 1} \frac{u_n}{n!} z^n$ ,  $F(z) = \sum_{n \geq 0} \frac{f_n}{n!} z^n$ , be the exponential generating functions of labeled: rooted trees, unrooted trees, and forests, respectively. Note that the sum defining  $F(z)$  starts at  $n = 0$  since we allow the empty forest in this discussion, and we set  $f_0 := 1$  accordingly. Since each rooted tree can be uniquely transformed, by removing the root, into a set made of smaller rooted trees, we have using the classical dictionary of symbolic combinatorics [FS09]:

$$T(z) = z \exp T(z). \quad (68)$$

Note that the closed formula  $t_n = n^{n-1}$  follows directly from the last equation and the Lagrange inversion formula.

In order to get an equation for the series  $U(z)$ , we note that any tree  $\mathbf{t}$  satisfies the equation  $1 = \#\text{vertices}(\mathbf{t}) - \#\text{edges}(\mathbf{t})$ . By summing this equation over all trees of given size, and observing that the generating function of trees rooted at an edge is given by  $\frac{1}{2}T(z)^2$ , we obtain the *disymmetry equation* (see [BLL98]):

$$U(z) = T(z) - \frac{1}{2}T(z)^2. \quad (69)$$

Finally, since any forest can be viewed as a set of *unrooted* trees, we have:

$$F(z) = \exp U(z). \quad (70)$$

The above equations express implicitly the functions  $T(z), U(z), F(z)$  in terms of the variable  $z$ . In order to transform these implicit equations into asymptotic estimates, we will use so called *transfer* theorems. These results (see [FS09, Chap VI]) enable one to deduce precise estimates on the coefficients of a generating function given an expansion at the main singularity of this function on its disk of convergence.

First, it follows by (68) and the implicit function theorem (see also the general theorems in [FS09, Chap VII]) that the function  $T(z)$  has radius of convergence  $e^{-1}$ , has a unique dominant singularity at  $z = e^{-1}$ , and has an expansion of the form:

$$T(z) = 1 + c\sqrt{1 - ze} + c'(1 - ze) + O((1 - ze)^{3/2}) \quad (71)$$

for some  $c, c' \neq 0$ , uniformly in a neighbourhood of  $z = e^{-1}$  slit along the half-line  $[e^{-1}, \infty)$ . Of course we could compute the value of  $c$  and  $c'$  by implicit differentiation, but this will not be needed. It follows from (69) that the series  $U(z)$  has, uniformly in the same domain, an expansion of the form:

$$U(z) = \frac{1}{2} + c''(1 - ze) + cc'(1 - ze)^{3/2} + O((1 - ze)^2), \quad (72)$$

and from (70) we obtain the expansion of the function  $F(z)$ :

$$F(z) = e^{1/2} + c'''(1 - ze) + e^{1/2}cc'(1 - ze)^{3/2} + O((1 - ze)^2). \quad (73)$$

Note that Equations (71) and (73) imply Lemma 6 stated in Section 2.4.

We are now ready to reprove the result of Rényi about the probability that a forest is connected [Rén59]. Applying a standard transfer theorem for analytic functions [FS09], (72) implies that the number of (unrooted) trees of size  $n$  satisfies, when  $n$  goes to infinity:

$$\frac{u_n}{n!} \sim \frac{cc'}{\Gamma(-3/2)} n^{-5/2} e^n.$$

Similarly, for  $f_n$ , (73) implies:

$$\frac{f_n}{n!} \sim \frac{cc'e^{1/2}}{\Gamma(-3/2)} n^{-5/2} e^n. \quad (74)$$

from which we get Renyi's result:

$$\frac{u_n}{f_n} \rightarrow e^{-1/2}.$$

More generally, fix a forest  $U \in \mathcal{F}$ , *possibly empty*, and let  $F^U(z)$  be the generating function of all forests  $\mathbf{f}$  such that  $Small(\mathbf{f}) \equiv U$ . Since there are  $\frac{|U|!}{\text{Aut}(U)}$  labeled forests isomorphic to  $U$ , we have:

$$F^U(z) = \frac{z^{|U|}}{\text{Aut}(U)} U(z) + P_U(z),$$

where  $P_U$  is a *polynomial* taking into account the cases of small forests containing  $U$  as a subforest but in which the largest component is among the components of  $U$ . The previous equation is also valid when  $U$  is empty (with  $\frac{z^{|U|}}{\text{Aut}(U)} := 1$ ).

From our previous estimates, the function  $F^U$  has an asymptotic expansion, uniformly in a neighbourhood of  $e^{-1}$  slit along  $[e^{-1}, \infty)$  of the form:

$$F^U(z) = c_4 + c_5(1 - ze) + \frac{e^{-|U|}}{\text{Aut}(U)} cc'(1 - ze)^{3/2} + O((1 - ze)^2). \quad (75)$$

(note that the unknown polynomial  $P_U$  can contribute to the constant or linear term, but not the singular term of exponent  $3/2$ , which fortunately is the only one we are interested in for the analysis). By standard transfer theorems, we get that when  $n$  goes to infinity the coefficient of  $z^n$  in this function is equivalent to:

$$\frac{f_n^U}{n!} \sim \frac{cc'e^{-|U|}}{\text{Aut}(U)\Gamma(-3/2)} n^{-5/2} e^n.$$

From (73), we thus obtain:

$$\frac{f_n^U}{f_n} \longrightarrow \frac{e^{-\frac{1}{2}-|U|}}{\text{Aut}(U)} = p_\infty(U),$$

which is precisely the first part of Theorem B. In order to see (as is also claimed in Theorem B) that  $p_\infty$  is a probability measure, we note that the sum over all (possibly empty) forests  $U \in \mathcal{F}$  given by:

$$\sum_{U \in \mathcal{F}} p_\infty(U) = e^{-1/2} \sum_{U \in \mathcal{F}} \frac{e^{-|U|}}{\text{Aut}(U)}$$

is equal to  $e^{-1/2}F(e^{-1})$  by definition of  $F(z)$  (recall that  $F(z)$  contains the contribution of the empty forest). Therefore the fact that  $p_\infty$  is a probability measure is equivalent to the equation  $e^{-1/2}F(e^{-1}) = 1$ , which is true from (73).

To prove the second part of Theorem B, we proceed once again by analysis of generating functions, together with the second moment method. Fix a rooted (unlabeled) tree  $T$ . Since  $T$  has  $\frac{|T|!}{\text{Aut}_r(T)}$  different labelings, the generating function of all labeled forests carrying a marked pendant copy of  $T$  is equal to:

$$\left( \frac{z^{|T|}}{\text{Aut}_r(T)} T(z) + Q_T(z) \right) F(z) \quad (76)$$

where the factor  $T(z)$  accounts for the tree attached to the other end of the edge from which the copy of  $T$  is pending, where  $F(z)$  accounts for the other components, and where as before  $Q_T(z)$  is a polynomial taking into account small number effects. The expansion of this quantity when  $z$  approaches  $e^{-1}$  is given, using previous expansions, by:

$$c_6 + \frac{c \cdot e^{-|T|+1/2}}{\text{Aut}_r(T)} \sqrt{1 - ze} + O(1 - ze).$$

By standard transfer theorems, when  $n$  goes to infinity the coefficient of  $z^n$  in this function, which is the number of forests of size  $n$  with a *marked* pendant copy of  $T$ , is equivalent to:

$$\frac{\Gamma(-1/2)}{\Gamma(-3/2)} \frac{c \cdot e^{-|T|+1/2}}{\text{Aut}_r(T)} n^{-3/2} e^n.$$

Dividing by  $f_n/n!$  and using (74) we get that the *average* number  $\chi_1(n)$  of pendant copies of  $T$  in a random forest of size  $n$  is asymptotic to:

$$\chi_1(n) \sim \frac{\Gamma(-1/2)}{e^{\Gamma(-3/2)}} \frac{e^{-|T|}}{\text{Aut}_r(T)} n = \frac{e^{-|T|}}{\text{Aut}_r(T)} n,$$

where we used that  $\frac{\Gamma(-1/2)}{e^{\Gamma(-3/2)}} = 1$  (which can be seen either by computing explicitly the expansion (71) or without computation by noting that  $t_n = n \cdot u_n$  and applying a transfer theorem to estimate  $t_n$  and  $u_n$  from (71) and (72)).

We have thus proved that the *average* number of pendant copies of  $T$  in a random forest of size  $n$  is equivalent to  $\frac{e^{-|T|}}{\text{Aut}_r(T)} n$ . To prove that this is in fact the typical value with high probability we

will compute the second moment. The generating function of forests with *two* marked copies of the tree  $T$  is of the form:

$$\left( \left( \frac{z^{|T|}}{\text{Aut}_r(T)} \right)^2 \left( \frac{zd}{dz} \right) T(z) + R_T(z)T(z) + S_T(z) \right) F(z) \quad (77)$$

where the first term is interpreted as a doubly rooted tree (hence  $(\frac{zd}{dz})T(z)$ ) to which two non-overlapping copies of  $T$  are attached, where the second term takes into account overlapping copies (here we do not need to know  $R_T(z)$  explicitly, but we know it is a polynomial) and the polynomial  $S_T(z)$  takes into account small cases for which there is ambiguity in the choice of the largest component; as before the factor  $F(z)$  takes into account the forest formed by other connected components. The expansion at  $z = e^{-1}$  of this function follows from differentiating (71) and is given by:

$$\frac{-c}{2} \left( \frac{e^{-|T|}}{\text{Aut}_r(T)} \right)^2 e^{1/2} (1 - ze)^{-1/2} + O(1).$$

Using a transfer theorem, the coefficient of  $z^n$  in this expression is asymptotic to:

$$\frac{-c}{2\Gamma(-1/2)} \left( \frac{e^{-|T|}}{\text{Aut}_r(T)} \right)^2 e^{1/2} n^{-1/2} e^n.$$

Dividing by  $f_n/n!$  and using (74), we see that the average squared number  $\chi_2(n)$  of pendant copies of  $T$  in a random forest of size  $n$  is asymptotic to:

$$\chi_2(n) \sim \frac{\Gamma(-3/2)}{2c\Gamma(-1/2)} \left( \frac{e^{-|T|}}{\text{Aut}_r(T)} \right)^2 n^2 = \left( \frac{e^{-|T|}}{\text{Aut}_r(T)} \right)^2 n^2,$$

where one may choose either previously mentioned method to compute the explicit value of  $c'$ . We have finally obtained that

$$\chi_2(n) \sim (\chi_1(n))^2,$$

so by the second moment method (Chebyshev's inequality) this completes the proof of the second part of Theorem B.

## A.2 More details on the example given in Remark 3

Let  $\tilde{\mathcal{F}}_n$  be the class of graphs defined in Remark 3, and write  $k_n := \lceil n^{2/3} \rceil$ . In this section we prove that  $\tilde{\mathcal{F}}_n$  is tight.

For  $i \geq k \geq 1$  let  $\tilde{a}_{i,k}$  be the number of connected graphs on  $[1..i]$  that induce a clique on  $[1..k]$ , and such that contracting this clique gives a tree. Thus the number of *connected* graphs in our class  $\tilde{\mathcal{F}}_n$  is, by definition, equal to  $\tilde{a}_{n,k_n}$ . Note that  $\tilde{a}_{i,k}$  equals to the number of rooted forests on  $[1..i]$  with  $k$  components rooted at  $1, 2, \dots, k$ . Thus  $\binom{i}{k} a_{i,k}$  is the number of rooted forests on  $[1..i]$  with  $k$  components and no condition on the location of the roots, which is classically equal to  $\binom{i-1}{i-k} i^{i-k}$ . We thus get:

$$\tilde{a}_{i,k} = k i^{i-k-1}.$$

Observe that:

$$\frac{\tilde{a}_{i,k}}{\tilde{a}_{i+1,k}} = \frac{i^{i-k-1}}{(i+1)^{i-k}} = \frac{1}{i} \left( 1 - \frac{1}{i+1} \right)^{i-k}. \quad (78)$$

The number  $g_n$  of all elements in the class  $\tilde{\mathcal{F}}_n$  is given by:

$$\frac{g_n}{(n - k_n)!} = \sum_{\substack{i+j=n \\ j \geq 0, i \geq k_n}} \frac{\tilde{a}_{i,k_n}}{(i - k_n)!} \times \frac{f_j}{j!} \quad (79)$$

where  $f_j$  counts unrooted forests, as before. In this sum,  $i$  is interpreted as the number of vertices in the connected component containing the clique, and we have distributed the labeling binomial  $\binom{n-k_n}{j}$  among factors.

From (73), given  $\epsilon$  we can choose  $\delta$  small enough and  $j_0$  large enough such that  $\sum_{j \leq j_0} f_j z^j \geq e^{1/2}(1 - \epsilon)$  for any  $z \geq e^{-1} - \delta$ . Also, given  $\delta$  and  $j_0$ , for  $n$  large enough, we have from (78) that for any  $i$  larger than  $n - j_0$ :

$$\frac{\tilde{a}_{i,k_n}/(i - k_n)!}{\tilde{a}_{i+1,k_n}/(i + 1 - k_n)!} \geq e^{-1} - \delta.$$

We can now lower bound the sum (79) by keeping the contribution of relatively small values of  $j$ . More precisely, for  $n$  large enough, we have:

$$\begin{aligned} (79) &\geq \sum_{j \leq j_0} \frac{\tilde{a}_{n-j,k_n}}{(n - j - k_n)!} \frac{f_j}{j!} \\ &\geq \frac{\tilde{a}_{n,k_n}}{(n - k_n)!} \sum_{j \leq j_0} (e^{-1} - \delta)^j \frac{f_j}{j!} \\ &\geq \frac{\tilde{a}_{n,k_n}}{(n - k_n)!} e^{1/2}(1 - \epsilon). \end{aligned}$$

Given  $\zeta$ , consider  $\epsilon = \zeta/2$ . If  $n$  is large enough and  $\tilde{F}_n$  is a uniform random graph in  $\tilde{\mathcal{F}}_n$ , we thus have

$$\Pr(\tilde{F}_n \text{ is connected}) = \frac{\tilde{a}_{n,k_n}}{g_n} \leq e^{-1/2}(1 - \epsilon)^{-1} \leq (1 + \zeta)e^{-1/2}.$$

Since this is true for every  $\zeta$ , the class  $\tilde{\mathcal{F}}$  is tight.